B2

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2001 (26.04.2001)

PCT

(10) International Publication Number
WO 01/29215 A2

(51) International Patent Classification[7]: C12N 15/12, C07K 14/705, 16/28, A61K 38/17, 31/70, 39/395, C12Q 1/68, G01N 33/53

(21) International Application Number: PCT/IL00/00664

(22) International Filing Date: 19 October 2000 (19.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
132459 19 October 1999 (19.10.1999) IL
134932 7 March 2000 (07.03.2000) IL

(71) Applicant *(for all designated States except US)*: **COMPUGEN LTD.** [IL/IL]; Pinchas Rosen St. 72, 69512 Tel Aviv (IL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SAVITZKY, Kinneret** [IL/IL]; Metodela St. 44, 69548 Tel Aviv (IL). **BERNSTEIN, Jeanne** [IL/IL]; Harimon St. 23, 40300 Kfar Yona (IL). **DAVID, Anat** [IL/IL]; Habrosh St. 80, 30500 Binyamina (IL).

(74) Agent: **REINHOLD COHN AND PARTNERS**; P.O.B. 4060, 61040 Tel Aviv (IL).

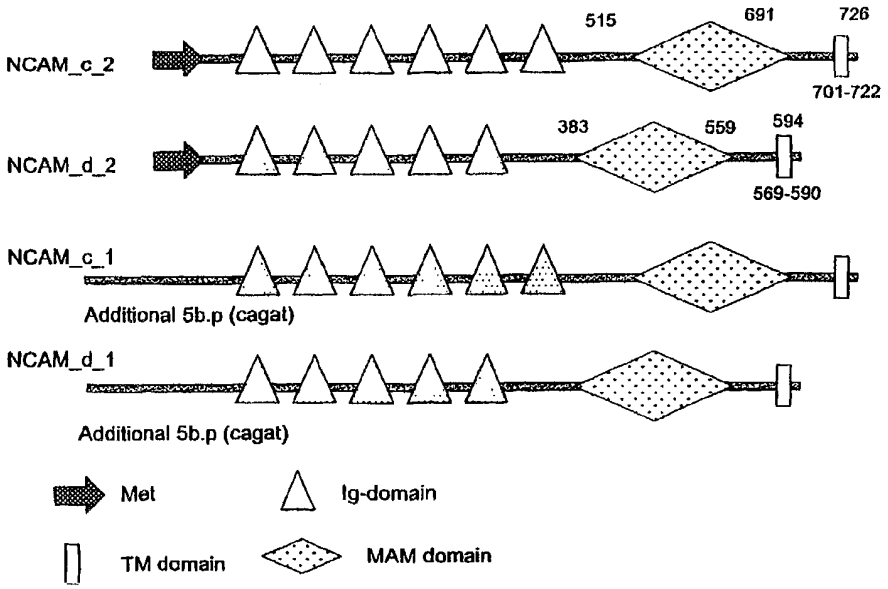
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

[Continued on next page]

(54) Title: N-CAM HOMOLOGS



(57) Abstract: The invention concerns several nucleic acid and amino acid sequences which are homologs of the neuronal cell adhesion molecule. These sequences are splice variants of each other.

WO 01/29215 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

N-CAM HOMOLOGS

FIELD OF THE INVENTION

The present invention concerns novel nucleic acid sequences, vectors and host cells containing them, amino acid sequences encoded by said sequences, and antibodies reactive with said amino acid sequences, as well as pharmaceutical compositions comprising any of the above. The present invention further concerns methods for screening for candidate activator or deactivators utilizing said amino acid sequences, as well as therapeutical and diagnostic utilization of said novel sequences.

BACKGROUND OF THE INVENTION

Cell adhesion molecules (CAMs) are transmembrane glycoproteins with extracellular binding domains and cytoplasmic functional domains. Ligand binding to the extracellular domain initiates intracellular events through the cytoplastmic functional domain. These in turn cause a major behavioral and functional change in the cells. For cells to interact with each other or with their matrix, two complementary molecules are required, the adhesion molecule and its ligand, one on each side of the adhesion site.

The immunoglobulin-like super family of adhesion molecules is a large and diverse family of molecules so named because they have one or more immunoglobulin-like domains. Included within the group are molecules concerned with antigen recognition by an adhesion to lymphocytes. These include CD3, CD4 and CD8 which together recognize complexes of antigen peptide and the major histocompatibility complex on other cells, and lymphocyte function related antigens such as CD2 and others of the LFA group of molecules.

Other important molecular subgroups in this super family are the intercellular adhesion molecules (ICAM) which are widely expressed on epithelial and endothelial cells, nervous system adhesion molecules, e.g. neural cell adhesion molecules (NICAM), and molecules including L1 and TAG involved in organization and function of nerves.

Cell adhesion molecules are responsible for more than just adhesion of cells to one another and to their insoluble matrices. Additional functions include the ordering of cell sorting, migration and differentiation; organization of cell motility via the cytoskeleton; regulation of inter and intracellular signaling; and control gene transcription.

Cell adhesion molecules (CAMs) are a subset of the immunoglobulin super family found in the nervous system of both vertebrates and invertebrates. They are usually surface membrane proteins with multiple Ig domains in their N-termini followed by several fibronectin-type III repeats and either a transmembrane intracellular domain or a glycophosphatidyl inositol-linked membrane anchor at the C-terminus.

The neuronal cell adhesion molecule (NCAM) is a member of the immunoglobulin super family and is strongly expressed in the nervous system. CAM is found in three major forms of which two – NCAM-140 and NCAM-180 are transmembrane proteins, while the third – NCAM-120 is attached to the membrane via a glycosylphosphatidyl inositol anchor. In addition, soluble NCAM forms exist in the brain, cerebrospinal fluid and in the plasma. NCAM mediates cell adhesion through homophilic as well as heterophilic activities. Following NCAM binding, transmembrane signalling is believed to be achieved resulting in increased intracellular calcium.

By mediating cell adhesion to other cells and to the extracellular matrix, and by activating intracellular signal pathways, NCAM influences cell migration, neurite extension and possibly also formation of synapses in the brain. From some studies on NCAM knock-out mice, it has been deduced that NCAM is crucial for the formation of the olfactory bulb and the mossy fiber system in the hippocampus.

In addition, NCAM is important for neuronal plasticity in the adult brain associated with learning and regeneration.

Several mutations in the L1 CAM are associated with the wide spectrum of neurologic abnormalities and mental retardation. This spectrum includes X-linked hydrocephalus, MASA syndrome, X-linked complicated spastic paraplegia type 1 and X-linked agenesis of the corpus callosum. These four diseases were initially described as distinct clinical entities with an overlapping clinical spectrum, but can now be lumped into one syndrome caused by mutations in the L1 gene. The main clinical features of this spectrum are Corpus callosum hypoplasia, mental retardation, adducted thumbs, spastic paraplegia and hydrocephalus, which has led to the acronym CRASH syndrome.

This protein may be a secreted adhesion-like molecule with anti-protease activity. A mutated or partially deleted sequence results in Kallmann Syndrome which is a genetic disorder that associates hypogonadotropic hypogonadism and anosmia. In this disease, the normal embryonic migration of GNRH-synthesizing neurons from the olfactory placodes to the hypothalamic region as well as the axonal extension of olfactory neurons towards the forebrain are impaired.

It has been found that NCAM is expressed also in a number of different tissues and cell types, beyond neurons, such as muscles and endocrine-originating tumors, and can be detected in sera of patients with small-cell lung cancer.

In tumor, the role of NCAM is attributed to tumor invasion and formation of metastatis. Malignant plasma cells and a subset of plasma cells from patients with monoclonal gammopathy exhibit surface expression of NCAM whereas normal plasma cells do not express NCAM.

An isoform of NCAM which is rich in polysialic acids and characteristic for embryonal NCAM (eNCAM) has been shown to be elevated in sera of patients with multiple myeloma using a chemiluminescence immunoassay.

GLOSSARY

In the following description and claims use will be made, at times, with a variety of terms, and the meaning of such terms as they should be construed in accordance with the invention is as follows:

"Cell adhesion molecules homolog (CAM-H) nucleic acid sequence" – the sequence shown in SEQ ID NO: 1 or 7, sequences having at least 70% identity to said sequence and fragments of the above sequences being 20 b.p. long. These sequences contain an N-terminal which has a homology to neural cell adhesion molecules containing several repeats of the Ig-like C2-domain which are immunoglobulin-like domains involved in protein-protein and protein-ligand interactions. In particular, the N-terminal part of the novel sequences of the invention are homologues to Contactin/F3-subgroup adhesion molecules; NB-2, Tag-1 (Axonin-1), Big-1, Big-2, and Contactin.

The novel sequences of the invention contain in their C-terminal neuropilin-like MAM domain. Neuropilin is a calcium-independent cell adhesion molecule that functions during the formation of certain neuronal circuits. This protein binds to Semaphorin III and to the VEGF165 isoform of VEGF. The MAM domain has been recognized as the extracellular region of functionally diverse proteins. All the protein have a modular receptor-like architecture consisting of a signal peptide, followed by a large N-terminal extracellular domain, a single transmembrane region and an intracellular region.

SEQ ID NOS: 1 to 6 are from Homo sapiens and SEQ ID NO: 7 is from mouse.

SEQ ID NO: 4 is in fact identical to SEQ ID NO: 1 – and was inserted under a different name for convenience reasons. SEQ ID NO: 5 is an update of SEQ ID NO: 2 (i.e. a sequence where 3 nucleotides have been altered).

SEQ ID NO: 1 (and 4), 2 (and 5), 3 and 6 are in fact all splice variants of each other.

The term CAM-H does not necessarily signify that the protein coded by the above sequences has the same or even similar physiological activities to known

CAM molecules and merely indicates that it shows sequence homology with two CAMs.

"CAM homology product (CAM-H product) – also referred at times as the "CAM-H protein" or ***"CAM-H polypeptide"*** – is an amino acid coded by the nucleic acid sequences of SEQ ID NO: 1 to SEQ ID NO: 7. The amino acid sequence may be a peptide, a protein, as well as peptides or proteins having *chemically modified* amino acids (see below) such as a glycopeptide or glycoprotein. An example of an CAM-H product is shown in SEQ ID NO: 8 to SEQ ID NO: 13 (coded by the amino acid sequences of SEQ ID NO: 1 to SEQ ID NO: 7 respectively). The term also includes analogues of said sequences in which one or more amino acids has been added, deleted, *substituted* (see below) or *chemically modified* (see below) as well as fragments of this sequence having at least 10 amino acids.

"Nucleic acid sequence" – a sequence composed of DNA nucleotides, RNA nucleotides or a combination of both types and may includes natural nucleotides, chemically modified nucleotides and synthetic nucleotides.

"Amino acid sequence" – a sequence composed of any one of the 20 naturally appearing amino acids, amino acids which have been *chemically modified* (see below), or composed of synthetic amino acids.

"Fragment of CAM-H product" - a polypeptide which has an amino acid sequence which is the same as part of but not all of the amino acid sequence of the CAM-H product.

"Fragments of CAM-H nucleic acid sequence" a continuous portion, preferably of about 20 nucleic acid sequences of the CAM-H nucleic acid sequence.

"Conservative substitution" - refers to the substitution of an amino acid in one class by an amino acid of the same class, where a class is defined by common physicochemical amino acid side chain properties and high substitution frequencies in homologous proteins found in nature, as determined, for example, by a standard Dayhoff frequency exchange matrix or BLOSUM matrix. [Six general classes of amino acid side chains have been categorized and include: Class I (Cys); Class II (Ser, Thr, Pro, Ala, Gly); Class III (Asn, Asp, Gln, Glu); Class IV (His, Arg, Lys); Class V (Ile, Leu, Val, Met); and Class VI (Phe, Tyr, Trp). For example, substitution of an Asp for another class III residue such as Asn, Gln, or Glu, is a conservative substitution.

"Non-conservative substitution" - refers to the substitution of an amino acid in one class with an amino acid from another class; for example, substitution of an Ala, a class II residue, with a class III residue such as Asp, Asn, Glu, or Gln.

"Chemically modified" - when referring to the product of the invention, means a product (protein) where at least one of its amino acid resides is modified either by natural processes, such as processing or other post-translational modifications, or by chemical modification techniques which are well known in the art. Among the numerous known modifications typical, but not exclusive examples include: acetylation, acylation, amidation, ADP-ribosylation, glycosylation, GPI anchor formation, covalent attachment of a lipid or lipid derivative, methylation, myristlyation, pegylation, prenylation, phosphorylation, ubiqutination, or any similar process.

"Biologically active" - refers to the CAM-H product which has physiological, regulatory or biochemical functions on the same target sites which naturally occurring CAM-H effects, for example can bind to the same ligands as the CAM, for example ligands present in extracellular matrixes; can interact with the same ligands present on the surface cells with which CAMs interact etc.

"Immunologically active" defines the capability of a natural, recombinant or synthetic CAM-H product, or any fragment thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies. Thus, for example, a biologically active fragment of CAM-H product denotes a fragment which retains some or all of the immunological properties of the CAM-H product, e.g can bind specific anti-CAM-H product antibodies or which can elicit an immune response which will generate such antibodies or cause proliferation of specific immune cells which produce CAM-H.

"Optimal alignment" - is defined as an alignment giving the highest percent identity score. Such alignment can be performed using a variety of commercially available sequence analysis programs, such as the local alignment program LALIGN using a ktup of 1, default parameters and the default PAM. A preferred alignment is the one performed using the CLUSTAL-W program from MacVector (TM), operated with an open gap penalty of 10.0, an extended gap penalty of 0.1, and a BLOSUM similarity matrix. If a gap needs to be inserted into a first sequence to optimally align it with a second sequence, the percent identity is calculated using only the residues that are paired with a corresponding amino acid residue (i.e., the calculation does not consider residues in the second sequences that are in the *"gap"* of the first sequence).

"Having at least X% identity" - with respect to two amino acid or nucleic acid sequence sequences, refers to the percentage of residues that are identical in the two sequences when the sequences are optimally aligned. Thus, 70% amino acid sequence identity means that 70% of the amino acids in two or more optimally aligned polypeptide sequences are identical.

"Isolated nucleic acid molecule having an CAM-H nucleic acid sequence" - is a nucleic acid molecule that includes the coding CAM-H nucleic acid sequence.

Said isolated nucleic acid molecule may include the CAM-H nucleic acid sequence as an independent insert; may include the CAM-H nucleic acid sequence fused to an additional coding sequences, encoding together a fusion protein in which the CAM-H coding sequence is the dominant coding sequence (for example, the additional coding sequence may code for a signal peptide); the CAM-H nucleic acid sequence may be in combination with non-coding sequences, e.g., introns or control elements, such as promoter and terminator elements or 5' and/or 3' untranslated regions, effective for expression of the coding sequence in a suitable host; or may be a vector in which the CAM-H protein coding sequence is a heterologous.

"Expression vector" - refers to vectors that have the ability to incorporate and express heterologous DNA fragments in a foreign cell. Many prokaryotic and eukaryotic expression vectors are known and/or commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.

"Deletion" - is a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent.

"Insertion" or "addition" - is that change in a nucleotide or amino acid sequence which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as compared to the naturally occurring sequence.

"Substitution" - replacement of one or more nucleotides or amino acids by different nucleotides or amino acids, respectively. As regards amino acid sequences the substitution may be conservative or non- conservative.

"Antibody" – refers to IgG, IgM, IgD, IgA, and IgG antibody. The definition includes polyclonal antibodies or monoclonal antibodies. This term refers to

whole antibodies or fragments of the antibodies comprising the antigen-binding domain of the anti-CAM-H product antibodies, e.g. antibodies without the Fc portion, single chain antibodies, fragments consisting of essentially only the variable, antigen-binding domain of the antibody, etc. The term *"antibody"* may also refer to antibodies against nucleic acid sequences obtained by cDNA vaccination.

"Activator" - as used herein, refers to a molecule which mimics the effect of the natural CAM-H product or at times even increases or prolongs the duration of the biological activity of said product, as compared to that induced by the natural product. The mechanism may be by binding to the same moiety to which native CAM-H binds (for example on the extracellular matrix or to a protein present on the membrane of other cells) thus mimicking the activity of CAM-H; by prolonging the lifetime of the CAM-H, (for example by decrease of the rate of its degradation), by increasing the affinity of CAM-H to moieties which it binds (such as the extracellular moieties) etc. Activators may be polypeptides, nucleic acids, carbohydrates, lipids, or derivatives thereof, or any other molecules which can bind to and activate the CAM-H product.

"Deactivator" or ***("Inhibitor")*** - refers to a molecule which modulates the activity of the CAM-H product in an opposite manner to that of the activator, by decreasing or shortening the duration of the biological activity of the CAM-H product. This may be done by blocking the binding of the CAM-H to the moiety for example to the ligand to which it binds by competitive or non-competitive inhibition, by causing rapid degradation of the CAM-H, etc. by inhibiting association of the CAM-H with the effectors which regulate its expression. Deactivators may be polypeptides, nucleic acids, carbohydrates, lipids, or derivatives thereof, or any other molecules which bind to and modulate the activity of said product.

"Treating a disease" - refers to administering a therapeutic substance effective to ameliorate symptoms associated with a disease, to lessen the severity or cure the disease, or to prevent the disease from occurring.

"Detection" – refers to a method of detection of a disease. This term may refer to detection of a predisposition to a disease.

"Probe" – the CAM-H nucleic acid sequence, or a sequence complementary therewith, when used to detect presence of other similar sequences in a sample. The detection is carried out by identification of hybridization complexes between the probe and the assayed sequence. The probe may be attached to a solid support or to a detectable label.

SUMMARY OF THE INVENTION

The present invention is based on the finding of novel cell adhesion molecules containing an N-terminal which has Ig-like C2 type repeats, and a C-terminal which contains a neuropilin-like MAM domain.

Thus the present invention provides by its first aspect, a novel isolated nucleic acid molecule comprising or consisting of the sequence of SEQ ID NO: 1 to SEQ ID NO: 7, fragments of said sequence having at least 20 nucleic acids, or a molecule comprising a sequence having at least 70%, preferably 80%, and most preferably 90% identity to SEQ ID NO:1 to SEQ ID NO: 7.

The present invention further provides a protein or polypeptide comprising or consisting of an amino acid sequence encoded by any of the above nucleic acid sequences, termed herein *"CAM-H product"*, for example, an amino acid sequence having the sequence as depicted in SEQ ID NO: 8 to 13, fragments of the above amino acid sequence having a length of at least 10 amino acids, as well as homologs of the amino acid sequences of SEQ ID NO: 8 to 13 in which one or more of the amino acid residues has been substituted (by conservative or non-conservative substitution) added, deleted, or chemically modified.

The present invention further provides nucleic acid molecule comprising or consisting of a sequence which encodes the above amino acid sequences, (including the fragments and analogs of the amino acid sequences). Due to the degenerative nature of the genetic code, a plurality of alternative nucleic acid sequences, beyond those depicted in SEQ ID NO:1 to SEQ ID NO: 7, can code for the amino acid sequence of the invention. Those alternative nucleic acid sequences which code for the same amino acid sequences codes by the sequences of SEQ ID NO: 1 to SEQ ID NO: 7 are also an aspect of the of the present invention.

The present invention further provides expression vectors and cloning vectors comprising any of the above nucleic acid sequences, as well as host cells transfected by said vectors.

The present invention still further provides pharmaceutical compositions comprising, as an active ingredient, said nucleic acid molecules, said expression vectors, or said protein or polypeptide.

These pharmaceutical compositions are suitable for the treatment of diseases and pathological conditions, which can be ameliorated, cured or prevented by raising the level of the CAM-H product. Typically these are diseases which are manifested by non-normal levels of various CAMs, which are usually lower than normal levels of CAM, or alternatively, diseases in which the level of CAM is normal, but a therapeutically beneficial effect may be achieved by raising the level of CAM to a higher than normal level. Usually, these diseases are concerned with interaction of cells at the extracellular matrix, or interaction with other cells which can lead to a plurality of diseases and pathological conditions, which are specified hereinbelow.

By this preferred embodiment the pharmaceutical compositions of the invention may stimulate growth and regeneration of nerve cell axons in order to compensate for defects caused by genetic, inflammatory, infectious, neurogenerative or trauma causes, for example to target neurite regrowth to target muscles and nerve cells after nerve injury.

In addition, the composition may be used for the treatment of a plurality of CAM-involved diseases such as: inflammatory diseases, autoimmune diseases (Crohn's disease, colitis, rheumatoid arthritis), graft vs. host and host vs. graft diseases, for the treatment of multiple sclerosis, diabetes, atherosclerosis, various types of cancer, for the treatment of injuries caused by head trauma as well as for the treatment of various viral disease and the treatment of various respiratory diseases.

By a second aspect, the present invention provides a nucleic acid molecule comprising or consisting of a non-coding sequence which is complementary to that of SEQ ID NO: 1 to SEQ ID NO: 7, or complementary to a sequence having at least 70% identity to said sequence or a fragment of said two sequences. The complementary sequence may be a DNA sequence which hybridizes with SEQ ID NO: 1 to SEQ ID NO: 7, or hybridizes to a portion of that sequence having a length sufficient to inhibit the transcription of the complementary sequence. The complementary sequence may be a DNA sequence which can be transcribed into an mRNA being an antisense to the mRNA transcribed from SEQ ID NO: 1 to SEQ ID NO: 7 or into an mRNA which is an antisense to a fragment of the mRNA transcribed from SEQ ID NO: 1 to SEQ ID NO: 7 which has a length sufficient to hybridize with the mRNA transcribed from SEQ ID NO: 1 to SEQ ID NO: 7, so as to inhibit its translation. The complementary sequence may also be the mRNA or the fragment of the mRNA itself.

The nucleic acid sequences of the second aspect of the invention may be used for therapeutic or diagnostic applications for example for detection of the expression of CAM-H in various tissues, such as neuronal tissue, endothelial or epithelial tissue, tissues connected to the immune system, tissue obtained from tumors as well as body fluids such as cerebrospinal fluid, plasma and blood. Said detection may be indicative of a plurality of diseases, and pathological conditions stemming from genetic, inflammatory, infectious, degenerative or trauma causes. The diseases which can be detected are the same diseases mentioned above in connection with the treatment by the therapeutical compositions of the invention

above. In addition, the detection may be indicative to the presence of tumors in general, and in particular to tumors from neuro-endocrine origin basis. In particular the level of any one of the CAM-H of the invention may be indicative of the invasiveness of the tumor indicating its stage of malignancy, which determination may be important for the prognosis of the patient, and for determining his/her best therapeutical modality.

The present invention also provides expression vectors comprising any one of the above defined complementary nucleic acid sequences and host cells transfected with said nucleic acid sequences or vectors, being complementary to those specified in the first aspect of the invention.

The invention also provides anti-CAM-H product antibodies, namely antibodies directed against the CAM-H product which specifically bind to said CAM-H product as well as antibodies which can be obtained through cDNA vaccination. Said antibodies are useful both for diagnostic and therapeutic purposes. For the diagnostic purposes they may be indicative of neuronal diseases and pathological conditions (such as for example inflammatory and autoimmune diseases, diseases of the respiratory and vascular tract, and indicative of tumors of neuro-endocrine origin and for the stage of malignancy as indicated above. In addition said antibody may be as an active ingredient in a pharmaceutical composition as will be explained below.

The present invention also provides pharmaceutical compositions comprising, as an active ingredient, the nucleic acid molecules which comprise or consist of said complementary sequences, or of a vector comprising said complementary sequences. The pharmaceutical composition thus provides pharmaceutical compositions comprising, as an active ingredient, said anti-CAM-H product antibodies.

The pharmaceutical compositions comprising said anti-CAM-H product antibodies or the nucleic acid molecule comprising said complementary sequence, are suitable for the treatment of diseases and pathological conditions where a therapeutically beneficial effect may be achieved by neutralizing the CAM-H or decreasing the amount of the CAM-H product or blocking its binding to its target

(for example the ligand to which it binds on cells), for example, by the neutralizing effect of the antibodies, or by the decrease of the effect of the antisense mRNA in decreasing expression level of the CAM-H product. Mostly these diseases are manifested by a higher than normal level of the CAM-H of the invention, or by normal level of CAMs, however the disease may be ameliorated or a beneficial effect may be evident by decreasing said level. Examples of such diseases are for example tumors, and in particular of a neuro-endocrine origin, in which the invasiveness of the tumor is dependent on interactions between the CAM-H of the invention and basal membranes or extracellular matrixes, diseases involving the immune system as well as endothelial and epithelial membranes, as well as other diseases of neural origin involving cell-adhesion molecules.

According to the third aspect of the invention the present invention provides methods for detecting the level of the transcript (mRNA) of said CAM-H product in a body fluid sample, plasma, cerebrospinal fluid, or in a specific tissue sample, for example by use of probes comprising or consisting of said coding sequences; as well as methods for detecting levels of expression of said product in tissue, e.g. by the use of antibodies capable of specifically reacting with the above amino acid sequences. Detection of the level of the expression of the CAM-H of the invention may be indicative of a plurality of physiological or pathological conditions, as explained above.

The method, according to this latter aspect, for detection of a nucleic acid sequence which encodes the CAM-H product in a biological sample, comprises the steps of:

(a) providing a probe comprising at least one of the nucleic acid sequence defined above;

(b) contacting the biological sample with said probe under conditions allowing hybridization of nucleic acid sequences thereby enabling formation of hybridization complexes;

(c) detecting hybridization complexes, wherein the presence of the complex indicates the presence of nucleic acid sequence encoding the CAM-H product in the biological sample.

The amount of hybridization complexes may be determined and calibrated by comparing it to a calibration scale in order to determine the amount of the nucleic acid sequence which enables the CAM-H product in the sample. The level of each of the sequences SEQ ID NO: 1 to SEQ ID NO: 7 may be detected and either compared to each other, and said ratio may also be indicative to a plurality of pathological or physiological conditions as explained above.

By a preferred embodiment the probe is part of a nucleic acid chip used for detection purposes, i.e. the probe is a part of an array of probes each present in a known location on a solid support.

The nucleic acid sequence used in the above method may be a DNA sequence an RNA sequence, etc; it may be a coding or a sequence or a sequence complementary thereto (for respective detection of RNA transcripts or coding-DNA sequences). By quantization of the level of hybridization complexes and calibrating the quantified results it is possible also to detect the level of the transcript in the sample.

Methods for detecting mutations in the region coding for the CAM-H product are also provided, which may be methods carried-out in a binary fashion, namely merely detecting whether there is any mismatches between the normal CAM-H nucleic acid sequence and the one present in the sample, or carried-out by specifically detecting the nature and location of the mutation.

The present invention also concerns a method for detecting CAM-H product both for determining its presence, as well as its level or alterations in its level in a biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of the invention, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of CAM-H product in said biological sample.

The present invention also concerns a method for detecting anti-CAM-H antibodies in a biological sample comprising the steps of:

(a) contacting said biological sample with the product of the invention thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of anti-CAM-H antibody in said biological sample.

In many cases, diseases are detected not by detecting the presence of the protein (product) which caused the disease, but rather by detecting the presence in a biological sample (such as blood or serum) of antibodies against such a product. The method of detecting the presence of anti-CAM-H antibodies is intended to be used in such case.

The amount of the antibody-antigen complex can be quantitized, in order to determine the level of the CAM-H-product or the anti-CAM-H antibodies, as the case may be.

As explained above the level of any of the products of SEQ ID NO: 8 (of the novel homolog) may be compared to each level, and the ratio between the levels may be indicative to a plurality of physiological and pathological conditions as explained above. In addition, the indicative ratio may not be the ratio of the proteins themselves but rather the ratio of antibodies against the proteins. Furthermore, the ratio of the level of the mRNA transcripts of any one of SQ ID NO: 1 to SEQ ID NO: 7, and changes in said ratio may be indicative of a plurality of diseases or pathological conditions especially as detailed above.

By yet another aspect the invention also provides a method for identifying candidate compounds capable of modulating the activity of CAM-H product (being either activators or deactivators). The method includes:

(i) providing a protein or polypeptide comprising an amino acid sequence substantially as depicted in SEQ ID NO: 8 to SEQ ID NO: 13, or a fragment of such a sequence;

(ii) contacting a candidate compound with said amino acid sequence;

(iii) comprising the physiological effect of the amino acid sequence in the presence and absence of said candidate compound and selecting those compounds which show a significant effect on said physiological activity.

The activity of the amino acid which should be changed by the modulator (being either the activator or deactivator) may be for example the binding of the

CAM-H product to a ligand of an adhesion molecule present for example on the external cell surface. Any modulator which changes such an activity has a potential as serving as an actuator or deactivator.

The present invention also concerns compounds identified by the above methods described above, which compound may either be an activator of the CAM-H product or a deactivator thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

In order to understand the invention and to see how it may be carried out in practice, a preferred embodiment will now be described, by way of non-limiting example only, with reference to the accompanying drawings, in which:

Fig. 1 is the alignment of the CAM-H product of SEQ ID NO: 3 (NCAM_c_1), 4 (NCAM_c_2); 5 (NCAM_d_2) and 6 (NCAM_d_1);

Fig. 2 is an alignment of human (SEQ ID NO: 4) and mouse (SEQ ID NO: 7) sequences;

Fig. 3 is an alignment of amino acid sequences of SEQ ID NO: 8 and 9;

Fig. 4 is an alignment of SEQ ID NOS: 10, 11, 12, 13;

Fig. 5 shows gel of PCR fragment of the sequence of the invention in various tissues; and

Fig. 6 shows a schematic drawing of the amino acid sequence of SEQ ID NOS; 10, 11, 12, 13, showing the positions of Tm, Ig, and MAM domains.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

Exámple I: CAM-H - nucleic acid sequence

The nucleic acid sequences of the invention include nucleic acid sequences which encode CAM-H product and fragments and analogs thereof. The nucleic acid sequences may alternatively be sequences complementary to the above coding sequence, or to a region of said coding sequence. The length of the complementary sequence is sufficient to avoid the expression of the coding

sequence. The nucleic acid sequences may be in the form of RNA or in the form of DNA, and include messenger RNA, synthetic RNA and DNA, cDNA, and genomic DNA. The DNA may be double-stranded or single-stranded, and if single-stranded may be the coding strand or the non-coding (anti-sense, complementary) strand. The nucleic acid sequences may also both include dNTPs, rNTPs as well as non naturally occurring sequences. The sequence may also be a part of a hybrid between an amino acid sequence and a nucleic acid sequence.

In a general embodiment, the nucleic acid sequence has at least 70%, preferably 80% or 90% sequence identity with the sequences identified as SEQ ID NO: 1 to SEQ ID NO: 7.

The nucleic acid sequences may include the coding sequence by itself. By another alternative the coding region may be in combination with additional coding sequences, such as those coding for fusion protein or signal peptides, in combination with non-coding sequences, such as introns and control elements, promoter and terminator elements or 5' and/or 3' untranslated regions, effective for expression of the coding sequence in a suitable host, and/or in a vector or host environment in which the CAM-H nucleic acid sequence is introduced as a heterologous sequence.

The nucleic acid sequences of the present invention may also have the product coding sequence fused in-frame to a marker sequence which allows for purification of the CAM-H product. The marker sequence may be, for example, a hexahistidine tag to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., *et al. Cell* 37:767 (1984)).

Also included in the scope of the invention are fragments also referred to herein as oligonucleotides, typically having at least 20 bases, preferably 20-30 bases corresponding to a region of the coding-sequence nucleic acid sequence.

The fragments may be used as probes, primers, and when complementary also as antisense agents, and the like, according to known methods.

As indicated above, the nucleic acid sequence may be substantially a depicted in SEQ ID NO: 1 to SEQ ID NO: 7 or fragments thereof or sequences having at least 70%, preferably 70-80%, most preferably 90% identity to the above sequence. Alternatively, due to the degenerative nature of the genetic code, the sequence may be a sequence coding the amino acid sequence of SEQ ID NO: 8 to SEQ ID NO: 13, or fragments or analogs of said amino acid sequence.

A. Preparation of nucleic acid sequences

The nucleic acid sequences may be obtained by screening cDNA libraries using oligonucleotide probes which can hybridize to or PCR-amplify nucleic acid sequences which encode the CAM-H products disclosed above. cDNA libraries prepared from a variety of tissues are commercially available and procedures for screening and isolating cDNA clones are well-known to those of skill in the art. Such techniques are described in, for example, Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2nd Edition), Cold Spring Harbor Press, Plainview, N.Y. and Ausubel FM et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

The nucleic acid sequences may be extended to obtain upstream and downstream sequences such as promoters, regulatory elements, and 5' and 3' untranslated regions (UTRs). Extension of the available transcript sequence may be performed by numerous methods known to those of skill in the art, such as PCR or primer extension (Sambrook *et al., supra*), or by the RACE method using, for example, the Marathon RACE kit (Clontech, Cat. # K1802-1).

Alternatively, the technique of *"restriction-site"* PCR (Gobinda *et al. PCR Methods Applic.* 2:318-22, (1993)), which uses universal primers to retrieve flanking sequence adjacent a known locus, may be employed. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second

round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. *et al.*, *Nucleic Acids Res.* **16**:8186, (1988)). The primers may be designed using OLIGO(R) 4.06 Primer Analysis Software (1992; National Biosciences Inc, Plymouth, Minn.), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Capture PCR (Lagerstrom, M. *et al.*, *PCR Methods Applic.* **1**:111-19, (1991)) is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into a flanking part of the DNA molecule before PCR.

Another method which may be used to retrieve flanking sequences is that of Parker, J.D., *et al.*, *Nucleic Acids Res.*, **19**:3055-60, (1991)). Additionally, one can use PCR, nested primers and PromoterFinder™ libraries to *"walk in"* genomic DNA (PromoterFinder™; Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions. Preferred libraries for screening for full length cDNAs are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes.

A randomly primed library may be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region.

The nucleic acid sequences and oligonucleotides of the invention can also be prepared by solid-phase methods, according to known synthetic methods. Typically, fragments of up to about 100 bases are individually synthesized, then joined to form continuous sequences up to several hundred bases.

B. Use of CAM-H nucleic acid sequence for the production of CAM-H products

In accordance with the present invention, nucleic acid sequences specified above may be used as recombinant DNA molecules that direct the expression of CAM-H products.

As will be understood by those of skill in the art, it may be advantageous to produce CAM-H product-encoding nucleotide sequences possessing codons other than those which appear in SEQ ID NO: 1 to SEQ ID NO: 7 which are those which naturally occur in the human genome. Codons preferred by a particular prokaryotic or eukaryotic-Host (Murray, E. *et al. Nuc Acids Res.,* 17:477-508, (1989)) can be selected, for example, to increase the rate of CAM-H product expression or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, than transcripts produced from naturally occurring sequence.

The nucleic acid sequences of the present invention can be engineered in order to alter a CAM-H product coding sequence for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of the product. For example, alterations may be introduced using techniques which are well known in the art, e.g., site-directed mutagenesis, to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to produce splice variants, etc.

The present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation.

In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are also described in Sambrook, *et al., (supra)*.

The present invention also relates to host cells which are genetically engineered with vectors of the invention, and the production of the product of the invention by recombinant techniques. Host cells are genetically engineered (i.e., transduced, transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the expression of the CAM-H nucleic acid sequence. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art.

The nucleic acid sequences of the present invention may be included in any one of a variety of expression vectors for expressing a product. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host. The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and related sub-cloning procedures are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate transcription control sequence (promoter) to direct mRNA synthesis. Examples of such promoters include: LTR or SV40 promoter, the *E.coli lac* or *trp* promoter, the phage lambda *PL* promoter, and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation, and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E.coli.*

The vector containing the appropriate DNA sequence as described above, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein. Examples of appropriate expression hosts include: bacterial cells, such as *E.coli, Streptomyces, Salmonella typhimurium*; fungal cells, such as yeast; insect cells such as *Drosophila* and *Spodoptera* Sf9; animal cells such as CHO, COS, HEK 293 or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein. The invention is not limited by the host cells employed.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the CAM-H product. For example, when large quantities of CAM-H product are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, multifunctional *E.coli* cloning and expression vectors such as *Bluescript*(R) (Stratagene), in which the CAM-H polypeptide coding sequence may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced;

pIN vectors (Van Heeke & Schuster *J. Biol. Chem.* 264:5503-5509, (1989)); *pET* vectors (Novagen, Madison WI); and the like.

In the yeast *Saccharomyces cerevisiae* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used. For reviews, see Ausubel *et al. (supra)* and Grant *et al.*, (*Methods in Enzymology* 153:516-544, (1987)).

In cases where plant expression vectors are used, the expression of a sequence encoding CAM-H product may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of *CaMV* (Brisson *et al., Nature* 310:511-514. (1984)) may be used alone or in combination with the omega leader sequence from TMV (Takamatsu *et al., EMBO J.*, 6:307-311, (1987)). Alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi *et al., EMBO J.* 3:1671-1680, (1984); Broglie *et al., Science* 224:838-843, (1984)); or heat shock promoters (Winter J and Sinibaldi R.M., *Results Probl. Cell Differ.*, 17:85-105, (1991)) may be used. These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. For reviews of such techniques, see Hobbs S. or Murry L.E. (1992) in McGraw Hill Yearbook of Science and Technology, McGraw Hill, New York, N.Y., pp 191-196; or Weissbach and Weissbach (1988) *Methods for Plant Molecular Biology*, Academic Press, New York, N.Y., pp 421-463.

CAM-H product may also be expressed in an insect system. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The CAM-H product coding sequence may be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of CAM-H coding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect *S. frugiperda* cells or *Trichoplusia* larvae in which CAM-H protein is expressed (Smith *et al.*,

J. Virol. 46:584, (1983); Engelhard, E.K. *et al., Proc. Nat. Acad. Sci.* **91**:3224-7, (1994)).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a CAM-H product coding sequence may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing CAM-H protein in infected host cells (Logan and Shenk, *Proc. Natl. Acad. Sci.* **81**:3655-59, (1984). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be required for efficient translation of a CAM-H protein coding sequence. These signals include the ATG initiation codon and adjacent sequences. In cases where CAM-H product coding sequence, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (Scharf, D. *et al.,* (1994) *Results Probl. Cell Differ.,* **20**:125-62, (1994); Bittner et al., *Methods in Enzymol* **153**:516-544, (1987)).

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium

phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., and Battey, I. (1986) Basic Methods in Molecular Biology). Cell-free translation systems can also be employed to produce polypeptides using RNAs derived from the DNA constructs of the present invention.

A host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the protein include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a *"pre-pro"* form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 293, WI38, etc. have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express CAM-H product may be transformed using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler M., *et al.*, *Cell* **11**:223-32, (1977)) and adenine phosphoribosyltransferase (Lowy I., *et al.*, *Cell* **22**:817-23, (1980)) genes which can be employed in *tk-* or *aprt-* cells, respectively. Also, antimetabolite,

antibiotic or herbicide resistance can be used as the basis for selection; for example, *dhfr* which confers resistance to methotrexate (Wigler M., *et al., Proc. Natl. Acad. Sci.* 77:3567-70, (1980)); *npt*, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al., J. Mol. Biol.*, 150:1-14, (1981)) and *als* or *pat*, which confer resistance to CAM-Horsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman S.C. and R.C. Mulligan, *Proc. Natl. Acad. Sci.* 85:8047-51, (1988)). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate, GUS, and luciferase and its substrates, luciferin and ATP, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. *et. al., Methods Mol. Biol.*, 55:121-131, (1995)).

Host cells transformed with a nucleotide sequence encoding CAM-H product may be cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The product produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing nucleic acid sequences encoding CAM-H product can be designed with signal sequences which direct secretion of CAM-H product through a prokaryotic or eukaryotic cell membrane.

CAM-H product may also be expressed as a recombinant protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle, Wash.). The inclusion of a

protease-cleavable polypeptide linker sequence between the purification domain and CAM-H protein is useful to facilitate purification. One such expression vector provides for expression of a fusion protein compromising a CAM-H polypeptide fused to a polyhistidine region separated by an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography, as described in Porath, *et al.*, *Protein Expression and Purification*, 3:263-281, (1992)) while the enterokinase cleavage site provides a means for isolating CAM-H polypeptide from the fusion protein. *pGEX* vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to ligand-agarose beads (e.g., glutathione-agarose in the case of GST-fusions) followed by elution in the presence of free ligand.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, or other methods, which are well know to those skilled in the art.

The CAM-H products can be recovered and purified from recombinant cell cultures by any of a number of methods well known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high

performance liquid chromatography (HPLC) can be employed for final purification steps.

C. Diagnostic applications utilizing nucleic acid sequences

The nucleic acid sequences of the present invention may be used for a variety of diagnostic purposes. The nucleic acid sequences may be used to detect and quantitate expression of CAM-H in patient's cells, e.g. biopsied tissues, by detecting the presence of mRNA coding for CAM-H product. Alternatively, the assay may be used to detect soluble CAM-H in the serum or blood. This assay typically involves obtaining total mRNA from the tissue or serum and contacting the mRNA with a nucleic acid probe. The probe is a nucleic acid molecule of at least 20 nucleotides, preferably 20-30 nucleotides, capable of specifically hybridizing with a sequence included within the sequence of a nucleic acid molecule encoding CAM-H under hybridizing conditions, detecting the presence of mRNA hybridized to the probe, and thereby detecting the expression of CAM-H. This assay can be used to distinguish between absence, presence, and excess expression of CAM-H product and to monitor levels of CAM-H expression during therapeutic intervention.

The invention also contemplates the use of the nucleic acid sequences as a diagnostic for diseases resulting from inherited defective CAM-H sequences. These sequences can be detected by comparing the sequences of the defective (i.e., mutant) CAM-H coding region with that of a normal coding region. Association of the sequence coding for mutant CAM-H product with abnormal CAM-H product activity may be verified. In addition, sequences encoding mutant CAM-H products can be inserted into a suitable vector for expression in a functional assay system (e.g., colorimetric assay, complementation experiments in a CAM-H protein deficient strain of HEK293 cells) as yet another means to verify or identify mutations. Once mutant genes have been identified, one can then screen populations of interest for carriers of the mutant gene.

Individuals carrying mutations in the nucleic acid sequence of the present invention may be detected at the DNA level by a variety of techniques. Nucleic acids used for diagnosis may be obtained from a patient's cells, including but not limited to such as from blood, urine, saliva, placenta, tissue biopsy and autopsy material. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki, *et al.*, *Nature* 324:163-166, (1986)) prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid of the present invention can be used to identify and analyze mutations in the gene of the present invention. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype.

Point mutations can be identified by hybridizing amplified DNA to radiolabeled RNA of the invention or alternatively, radiolabeled antisense DNA sequences of the invention. Sequence changes at specific locations may also be revealed by nuclease protection assays, such RNase and S1 protection or the chemical cleavage method (e.g. Cotton, *et alProc. Natl. Acad. Sci.* USA, 85:4397-4401, (1985)), or by differences in melting temperatures. *"Molecular beacons"* (Kostrikis L.G. *et al.*, Science 279:1228-1229, (1998)), hairpin-shaped, single-stranded synthetic oligo- nucleotides containing probe sequences which are complementary to the nucleic acid of the present invention, may also be used to detect point mutations or other sequence changes as well as monitor expression levels of CAM-H product. Such diagnostics would be particularly useful for prenatal testing.

Another method for detecting mutations uses two DNA probes which are designed to hybridize to adjacent regions of a target, with abutting bases, where the region of known or suspected mutation(s) is at or near the abutting bases. The two probes may be joined at the abutting bases, e.g., in the presence of a ligase enzyme, but only if both probes are correctly base paired in the region of probe junction. The presence or absence of mutations is then detectable by the presence or absence of ligated probe.

Also suitable for detecting mutations in the CAM-H product coding sequence are oligonucleotide array methods based on sequencing by hybridization (SBH), as described, for example, in U.S. Patent No. 5,547,839. In a typical method, the DNA target analyte is hybridized with an array of oligonucleotides formed on a microchip. The sequence of the target can then be *"read"* from the pattern of target binding to the array.

D. Gene mapping utilizing nucleic acid sequences

The nucleic acid sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 20-30 bp) from the CAM-H cDNA. Computer analysis of the 3' untranslated region is used to rapidly select primers that do not span more than one exon in the genomic DNA, which would complicate the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids or using instead radiation hybrids are rapid procedures for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in*

situ hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60 bases. For a review of this technique, see Verma *et al., Human Chromosomes: a Manual of Basic Techniques*, (1988) Pergamon Press, New York.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in the OMIM database (Center for Medical Genetics, Johns Hopkins University, Baltimore, MD and National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD). The OMIM gene map presents the cytogenetic map location of disease genes and other expressed genes. The OMIM database provides information on diseases associated with the chromosomal location. Such associations include the results of linkage analysis mapped to this interval, and the correlation of translocations and other chromosomal aberrations in this area with the advent of various diseases associated with abnormal amounts or function of various CAM proteins.

E. Therapeutic applications of nucleic acid sequences

Nucleic acid sequences of the invention may also be used for therapeutic purposes. Turning first to the second aspect of the invention (i.e. inhibition of expression of CAM-H), expression of CAM-H product may be modulated through antisense technology, which controls gene expression through hybridization of complementary nucleic acid sequences, i.e. antisense DNA or RNA, to the control, 5' or regulatory regions of the gene encoding CAM-H product. For example, the 5' coding portion of the nucleic acid sequence sequence which codes for the product of the present invention is used to design an antisense oligonucleotide of from about 10 to 40 base pairs in length.

Oligonucleotides derived from the transcription CAM-Ht site, e.g. between positions -10 and +10 from the CAM-Ht site, are preferred. An antisense DNA oligonucleotide is designed to be complementary to a region of the nucleic acid sequence involved in transcription (Lee *et al., Nucl. Acids, Res.*, 6:3073, (1979); Cooney et al., Science 241:456, (1988); and Dervan *et al., Science* 251:1360, (1991)), thereby preventing transcription and the production of the CAM-H products. An antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into the CAM-H products (Okano *J. Neurochem.* 56:560, (1991)). The antisense constructs can be delivered to cells by procedures known in the art such that the antisense RNA or DNA may be expressed *in vivo*. The antisense may be antisense mRNA or DNA sequence capable of coding such antisense mRNA. The antisense mRNA or the DNA coding thereof can be complementary to the full sequence of nucleic acid sequences coding to the CAM-H protein or to a fragment of such a sequence which is sufficient to inhibit production of a protein product.

Turning now to the first aspect of the invention, i.e. expression of CAM-H, expression of CAM-H product may be increased by providing coding sequences for coding for said product under the control of suitable control elements ending its expression in the desired host.

The nucleic acid sequences of the invention may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically effective amount of the compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration.

The products of the invention as well as any activators and deactivators compounds (see below) which are polypeptides, may also be employed in accordance with the present invention by expression of such polypeptides *in vivo*, which is often referred to as *"gene therapy."* Cells from a patient may be engineered with a nucleic acid sequence (DNA or RNA) encoding a polypeptide

ex vivo, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, cells may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding a polypeptide of the present invention.

Similarly, cells may be engineered in vivo for expression of a polypeptide *in vivo* by procedures known in the art. As known in the art, a producer cell for producing a retroviral particle containing RNA encoding the polypeptide of the present invention may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a product of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention. For example, the expression vehicle for engineering cells may be other than a retrovirus, for example, an adenovirus which may be used to engineer cells *in vivo* after combination with a suitable delivery vehicle.

Retroviruses from which the retroviral plasmid vectors mentioned above may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the *PE501*, *PA317*, *psi-2*, *psi-AM*, *PA12*, *T19-14X*, *VT-19-17-H2*, *psi-CRE*, *psi-CRIP*, *GP+E-86*, *GP+envAm12*, and *DAN* cell lines as described in Miller (*Human Gene Therapy*, Vol. 1, pg. 5-14, (1990)). The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and $CaPO_4$ precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either in vitro or in vivo. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

The genes introduced into cells may be placed under the control of inducible promoters, such as the radiation-inducible Egr-1 promoter, (Maceri, H.J., *et al., Cancer Res.*, 56(19):4311 (1996)), to stimulate CAM-H production or antisense inhibition in response to radiation, eg., radiation therapy for treating tumors.

Example II. CAM-H product

The substantially purified CAM-H product of the invention has been defined above as the product coded from the nucleic acid sequence of the invention. Preferably the amino acid sequence is an amino acid sequence having at least 70%, preferably at least 80% or 90% identity to the sequence identified as SEQ ID NO: 8 to SEQ ID NO: 13. The protein or polypeptide may be in mature and/or modified form, also as defined above. Also contemplated are protein fragments having at least 10 contiguous amino acid residues, preferably at least 10-20 residues, derived from the CAM-H product.

The sequence variations are preferably those that are considered conserved substitutions, as defined above. Thus, for example, a protein with a sequence having at least 80% sequence identity with the protein identified in SEQ ID NO: 8 to SEQ ID NO: 13, preferably by utilizing conserved substitutions as defined above is also part of the invention. In a more specific embodiment, the protein has or contains the sequence identified as SEQ ID NO: 8 to SEQ ID NO: 13. The CAM-H product may be (i) one in which one or more of the amino

acid residues in a sequence listed above are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue), or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the CAM-H product is fused with another compound, such as a compound to increase the half-life of the protein (for example, polyethylene glycol (PEG)), or a moiety which serves as targeting means to direct the protein to its target tissue or target cell population (such as an antibody), or (iv) one in which additional amino acids are fused to the CAM-H product. Such fragments, variants and derivatives are deemed to be within the scope of those skilled in the art from the teachings herein.

A. Preparation of CAM-H product

Recombinant methods for producing and isolating the CAM-H product, and fragments of the protein are described above.

In addition to recombinant production, fragments and portions of CAM-H product may be produced by direct peptide synthesis using solid-phase techniques (cf. Stewart *et al.*, (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco; Merrifield J., *J. Am. Chem. Soc.*, 85:2149-2154, (1963)). *In vitro* peptide synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, Calif.) in accordance with the instructions provided by the manufacturer. Fragments of CAM-H product may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

B. Therapeutic uses and compositions utilizing the CAM-H product

The CAM-H product of the invention is generally useful in treating diseases and disorders which are characterized by a lower than normal level of CAM-H expression, and or diseases which can be cured or ameliorated by raising the level of the CAM-H product, even if the level is normal. Typically, there are

diseases where non-normal growth, regeneration or tumors of neurons is evidenced due to genetic, degenerative or injury causes.

Typically these diseases are in CAM-H products or fragments and may be administered by any of a number of routes and methods designed to provide a consistent and predictable concentration of compound at the target organ or tissue. The product-containing compositions may be administered alone or in combination with other agents, such as stabilizing compounds, and/or in combination with other pharmaceutical agents such as drugs or hormones.

CAM-H product-containing compositions may be administered by a number of routes including, but not limited to, oral, intravenous, intramuscular, transdermal, subcutaneous, topical, sublingual, or rectal means as well as by nasal application. CAM-H product-containing compositions may also be administered via liposomes. Such administration routes and appropriate formulations are generally known to those of skill in the art.

The product can be given via intravenous or intraperitoneal injection. Similarly, the product may be injected to other localized regions of the body. The product may also be administered via nasal insufflation. Enteral administration is also possible. For such administration, the product should be formulated into an appropriate capsule or elixir for oral administration, or into a suppository for rectal administration.

The foregoing exemplary administration modes will likely require that the product be formulated into an appropriate carrier, including ointments, gels, suppositories. Appropriate formulations are well known to persons skilled in the art.

Dosage of the product will vary, depending upon the potency and therapeutic index of the particular polypeptide selected.

A therapeutic composition for use in the treatment method can include the product in a sterile injectable solution, the polypeptide in an oral delivery vehicle, the product in an aerosol suitable for nasal administration, or the product in a nebulized form, all prepared according to well known methods. Such

compositions comprise a therapeutically effective amount of the compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The product of the invention may also be used to modulate endothelial differentiation and proliferation as well as to modulate apoptosis either *ex vivo* or *in vitro*, for example, in cell cultures.

Example III. Screening methods for activators and deactivators (inhibitors)

The present invention also includes an assay for identifying molecules, such as synthetic drugs, antibodies, peptides, or other molecules, which have a modulating effect on the activity of the CAM-H product, e.g. activators or deactivators of the CAM-H product of the present invention. Such an assay comprises the steps of providing an CAM-H product encoded by the nucleic acid sequences of the present invention and determining its physiological activity on the target in the presence and absence of one or more candidate molecules to determine the candidate molecules. Those molecules which are modulating effect on the activity of the CAM-H product are selected as likely candidates for activators and deactivators.

CAM-H product, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening therapeutic compounds in any of a variety of drug screening techniques. The fragment employed in such a test may be free in solution, affixed to a solid support, borne on a cell membrane or located intracellularly. The formation of binding complexes, between CAM-H product and the agent being tested, may be measured. Alternatively, the activator or deactivator may work by serving as agonist or antagonist, respectively, of the CAM-H receptor and their effect may be determined in connection with the receptor.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the CAM-H product is described in detail by Geysen in PCT Application WO

84/03564, published on Sep. 13, 1984. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with the full CAM-H product or with fragments of CAM-H product and washed. Bound CAM-H product is then detected by methods well known in the art. Substantially purified CAM-H product can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

Antibodies to the CAM-H product, as described in Example IV below, may also be used in screening assays according to methods well known in the art. For example, a *"sandwich"* assay may be performed, in which an anti-CAM-H antibody is affixed to a solid surface such as a microtiter plate and CAM-H product is added. Such an assay can be used to capture compounds which bind to the CAM-H product. Alternatively, such an assay may be used to measure the ability of compounds to influence with the binding of CAM-H product to the CAM-H receptor and then select those compounds which effect the binding.

Example IV. Anti-CAM-H antibodies

A. Synthesis

In still another aspect of the invention, the purified CAM-H product is used to produce anti-CAM-H antibodies which have diagnostic and therapeutic uses related to the activity, distribution, and expression of the CAM-H product, in particular therapeutic applications in modulating the effect of CAM-H on moieties to which it binds in the extracellular matrix.

Antibodies to CAM-H product may be generated by methods well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments and fragments produced by an Fab expression library. Antibodies, i.e., those which inhibit dimer formation, are especially preferred for therapeutic use.

A fragment CAM-H product for antibody induction does not require biological activity but have to feature immunological activity; however, the protein fragment or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids of the sequences specified in SEQ ID NO: 8 to SEQ ID NO: 13. Preferably they should mimic a portion of the amino acid sequence of the natural protein and may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of CAM-H protein amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule. Procedures well known in the art can be used for the production of antibodies to CAM-H product.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, etc may be immunized by injection with CAM-H product or any portion, fragment or oligopeptide which retains immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include but are not limited to Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are potentially useful human adjuvants.

Monoclonal antibodies to CAM-H protein may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Koehler and Milstein (*Nature* **256**:495-497, (1975)); the human B-cell hybridoma technique (Kosbor *et al., Immunol. Today* **4**:72, (1983); Cote *et al., Proc. Natl. Acad. Sci.* **80**:2026-2030, (1983)) and the EBV-hybridoma technique (Cole, *et al., Mol. Cell Biol.* 62:109-120, (1984)).

Techniques developed for the production of *"chimeric antibodies"*, the splicing of mouse antibody genes to human antibody genes to obtain a molecule

with appropriate antigen specificity and biological activity can also be used (Morrison *et al., Proc. Natl. Acad. Sci.* 81:6851-6855, (1984); Neuberger *et al., Nature* 312:604-608, (1984); Takeda *et al., Nature* 314:452-454, (1985)). Alternatively, techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778) can be adapted to produce single-chain antibodies specific for the CAM-H protein.

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi *et al.* (*Proc. Natl. Acad. Sci.* 86:3833-3837, 1989)), and Winter G and Milstein C., (*Nature* 349:293-299, (1991)).

Antibody fragments which contain specific binding sites for CAM-H protein may also be generated. For example, such fragments include, but are not limited to, the $F(ab')_2$ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse W.D. *et al., Science* 256:1275-1281, (1989)).

B. Diagnostic applications of antibodies

A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the formation of complexes between CAM-H product and its specific antibody and the measurement of complex formation. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two noninterfering epitopes on a specific CAM-H product is preferred, but a competitive binding assay may also be employed. These assays are described in Maddox D.E., *et al.,* (*J. Exp. Med.* 158:1211, (1983)).

Antibodies which specifically bind CAM-H product are useful for the diagnosis of conditions or diseases characterized by over or under expression of CAM-H. Alternatively, such antibodies may be used in assays to monitor patients being treated with CAM-H product, its activators, or its deactivators. Diagnostic assays for CAM-H protein include methods utilizing the antibody and a label to detect CAM-H product in human body fluids or extracts of cells or tissues. The products and antibodies of the present invention may be used with or without modification. Frequently, the proteins and antibodies will be labeled by joining them, either covalently or noncovalently, with a reporter molecule. A wide variety of reporter molecules are known in the art.

A variety of protocols for measuring CAM-H product, using either polyclonal or monoclonal antibodies specific for the respective protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescent activated cell sorting (FACS). As noted above, a two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on CAM-H product is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, *et al. (supra).* Such protocols provide a basis for diagnosing altered or abnormal levels of CAM-H product expression. Normal or standard values for CAM-H product expression are established by combining body or cell extracts taken from normal subjects, preferably human, with antibody to CAM-H product under conditions suitable for complex formation which are well known in the art. The amount of standard complex formation may be quantified by various methods, preferably by photometric methods. Then, standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by disease. Deviation between standard and subject values establishes the presence of disease state.

The antibody assays are useful to determine the level of CAM-H present in a body fluid sample, in order to determine whether it is being overexpressed or

underexpressed in the tissue, or as an indication of how CAM-H levels are responding to drug treatment.

Another alternative is to determine the presence and/or level of naturally occurring anti-CAM-H antibodies in a sample, such as blood or serum. Many times diseases are identified by detecting the presence or level of antibodies against a specific product. For the detection of such naturally occurring anti-CAM-H antibodies, the sample may be contacted with the product of the invention, for example as depicted in **SEQ ID NO: 3 or SEQ ID NO: 4**, or with an antigenic fragment thereof, and the presence or level of antibody-antigen complexes may be determined by methods well known in the art.

C. Therapeutic uses of antibodies

In addition to their diagnostic use the antibodies may have a therapeutical utility in blocking or decreasing the activity of the CAM-H product in pathological conditions where beneficial effect can be achieved by such a decrease.

The antibody employed is preferably a humanized monoclonal antibody, or a human Mab produced by known globulin-gene library methods. The antibody is administered typically as a sterile solution by IV injection, although other parenteral routes may be suitable. Typically, the antibody is administered in an amount between about 1-15 mg/kg body weight of the subject. Treatment is continued, e.g., with dosing every 1-7 days, until a therapeutic improvement is seen.

Although the invention has been described with reference to specific methods and embodiments, it is appreciated that various modifications and changes may be made without departing from the invention.

EXAMPLE V. Specific localizations of the sequence of the invention in brain tissue

RT-PCR

Five μg of total RNA were isolated from human tissues. The reaction was performed in a final volume of 20 μl and also contained Superscript II Reverse Transcriptase (Gibco/BRL, Gaithersburg, MD), 1xBuffer supplied by the manufacturer, 30 units of Rnasin (Promega, Medison, WI) and 10 pmol of oligoDT (Promega, Medison, WI).

Gene-specific primers were designed, antisense and sense PCR was carried out in a reaction mixture containing 1 μl of 5xBuffer (Boehringer-Mannheim, Germany), 4 μl of 2.5 mM dNTPs, 25 pmol of primers and 2.5 units Expand Long Template PCR System (Boehringer-Mannheim, Germany).

The cycling conditions were 94°C-3 min. Taq polymerase followed by 30 cycles of 94°C for 30 sec, 68°C-1 min and final extension at 68°C for 10 mins. The PCR reaction on PTC-225 (MJ Research, Inc).

The resulting sequences were separated on a gel and the results are shown in Fig. 5. As can be seen the transcript (from SEQ ID NO: 3) was identified in Brain Cerebellum and was not identified in tissues obtained from other tissues proving that the sequences of the invention are indeed unique to brain.

CLAIMS:

1. An isolated nucleic acid sequence selected from the group consisting of:

(i) the nucleic acid sequence depicted in SEQ ID NO: 1 to SEQ ID NO: 7;

(ii) nucleic acid sequences having at least 70% identity with the sequence of (i); and

(iii) fragments of (i) or (ii) of at least 20 b.p.

2. A nucleic acid sequence according to Claim 1(ii) wherein the nucleic acid sequences have at least 80% identity with the sequence of Claim 1(i).

3. A nucleic acid sequence according to Claim 2, wherein the nucleic acid sequences have at least 90% identity.

4. An isolated nucleic acid sequence complementary to the nucleic acid sequence of Claim 1.

5. An amino acid sequence selected from the group consisting of:

(i) an amino acid sequence coded by the isolated nucleic acid sequence of Claim 1;

(ii) fragments of the amino acid squence of (i) having at least 10 amino acids;

(iii) analogues of the amino acid sequences of (i) or (ii) in which one or more amino acids has been added, deleted, replaced or chemically modified without substantially altering the biological activity of the parent amino acid sequence.

6. An amino acid sequence according to Claim 5, as depicted in SEQ ID NO: 8 to SEQ ID NO: 13.

7. An isolated nucleic acid sequence coding for the amino acid sequence of Claim 5 or 6.

8. A purified antibody which binds specifically to the amino acid sequence of Claim 5 or 6.

9. An expression vector comprising the nucleic acid sequences of Claim 1 or 7 and control elements for the expression of the nucleic acid sequence in a suitable host.

10. An expression vector comprising the nucleic acid sequence of Claim 4, and control elements for the expression of the nucleic acid sequence in a suitable host.

11. A host cell transfected by the expression vector of Claim 9 or 10.

12. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and as an active ingredient an agent selected from the group consisting of:

(i) the expression vector of Claim 9; and

(ii) the amino acid sequence of Claim 5 or 6.

13. A pharmaceutical composition according to Claim 12, for treatment of diseases which can be ameliorated, cured or prevented by raising the level of the Cell Adhesion Molecule Homolog (CAM-H).

14. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and as an active ingredient an agent selected from the group consisting of:

(i) the nucleic acid sequence of Claim 4;

(ii) the expression vector of Claim 10; and

(iii) the purified antibody of Claim 8.

15. A pharmaceutical composition according to Claim 14, for treatment of diseases which can be ameliorated or cured by decreasing the level of the CAM-H product.

16. A method for detecting an CAM-H nucleic acid sequence in a biological sample, comprising the steps of:

(a) hybridizing to nucleic acid material of said biological sample a nucleic acid sequence of Claim 1 or 4; and

(b) detecting said hybridization complex;

wherein the presence of said hybridization complex correlates with the presence of an CAM-H nucleic acid sequence in the said biological sample.

17. A method according to Claim 16, wherein the nucleic acid material of said biological sample are mRNA transcripts.

18. A method according to Claim 16, where the nucleic acid sequence is present in a nucleic acid chip.

19. A method for identifying candidate compounds capable of binding to the CAM-H product and modulating its activity the method comprising:

(i) providing a protein or polypeptide comprising an amino acid sequence substantially as depicted in SEQ ID NO: 8 to SEQ ID NO: 13, or a fragment of such a sequence;

(ii) comparing the physiological effect of the CAM-H product in the absence and presence of said candidate compound and selecting those compounds which show a significant effect on said physiological activity.

20. A method according to Claim 19, wherein the compound is an activator and the measured effect is increase in the physiological activity.

21. A method according to Claim 19, wherein the compound is an deactivator and the effect is decrease in the physiological activity.

22. An activator of the amino acid sequence of Claim 5 or 6.

23. An deactivator of the amino acid sequence of Claims 5 or 6.

24. A method for detecting CAM-H-product in a biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of Claim 8, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of CAM-H product in said biological sample.

25. A method for detecting anti-CAM-H antibodies in a biological sample comprising the steps of:

(a) contacting with said biological sample the antibody of Claim 5 or 6, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of anti-CAM-H antibody in said biological sample.

```
                                                                  50
NCAM_c_1  GTTTTTTACC AACAAGGCAG AGGATCAGAG ATGTCTTGCT TCAGCTCCTC
NCAM_c_2  GTTTTTTACC AACAAGGCAG AGGATCAGAG ATGTCTTGCT TCAGCTCCTC
NCAM_d_2  GTTTTTTACC AACAAGGCAG AGGATCAGAG ATGTCTTGCT TCAGCTCCTC
NCAM_d_1  GTTTTTTACC AACAAGGCAG AGGATCAGAG ATGTCTTGCT TCAGCTCCTC

                                                                 100
NCAM_c_1  CCACGGTTCG TATTGTGCAC TCAGGCTTGG CCTGTAACAT TGAAGAGGAG
NCAM_c_2  CCACGGTTCG TATTGTGCAC TCAGGCTTGG CCTGTAACAT TGAAGAGGAG
NCAM_d_2  CCACGGTTCG TATTGTGCAC TCAGGCTTGG CCTGTAACAT TGAAGAGGAG
NCAM_d_1  CCACGGTTCG TATTGTGCAC TCAGGCTTGG CCTGTAACAT TGAAGAGGAG

                                                                 150
NCAM_c_1  CGCTACTCCG AAAGGGTCTA CACTATCCGG GAAGGAGAAA CCCTtGAGTT
NCAM_c_2  CGCTACTCCG AAAGGGTCTA CACTATCCGG GAAGGAGAAA CCCTtGAGTT
NCAM_d_2  CGCTACTCCG AAAGGGTCTA CACTATCcGG gAAgGAGaAA CCCTtGAGTT
NCAM_d_1  CGCTACTCCG AAAGGGTCTA CACTATCcGG gAAgGAGaAA CCCTtGAGTT

                                                                 200
NCAM_c_1  GACTTGTCTA GTcACAGGaC ATCCACGTCc AcagatCAGG TGGacCAAAA
NCAM_c_2  GACTTGTCTA GTcACAGGaC ATCCACGTCc A.....CAGG TGGacCAAAA
NCAM_d_2  GACTTGTCTA GTcACAGGaC ATCCACGTCc A.....CAGG TGGacCAAAA
NCAM_d_1  GACTTGTCTA GTcACAGGaC ATCCACGTCc AcagatCAGG TGGacCAAAA
```

Fig. 1

```
          201                                                 250
NCAM_c_1  CAgCAGgAAG TGCcTcTGaC AgATTCCAAG ACTCAAGTGT CTTCAATGAG
NCAM_c_2  CAgCAGgAAG TGCcTcTGaC AgATTCCAAG ACTCAAGTGT CTTCAATGAG
NCAM_d_2  CAgCAGgAAG TGCcTcTGaC AgATTCCAAG ACTCAAGTGT CTTCAATGAG
NCAM_d_1  CAgCAGgAAG TGCcTcTGaC AgATTCCAAG ACTCAAGTGT CTTCAATGAG

          251                                                 300
NCAM_c_1  aCTTTGAGGA TTACAAATAT TCAGCGACAC CAAGGAGGCc GGTATTaCTG
NCAM_c_2  aCTTTGAGGA TTACAAATAT TCAGCGACAC CAAGGAGGCc GGTATTaCTG
NCAM_d_2  aCTTTGAGGA TTACAAATAT TCAGCGACAC CAAGGAGGCc GGTATTaCTG
NCAM_d_1  aCTTTGAGGA TTACAAATAT TCAGCGACAC CAAGGAGGCc GGTATTaCTG

          301                                                 350
NCAM_c_1  TAAAGCAGAG AATGGcTTGG GGTCTCCAGC gATAAAGTCA ATCAGAGTGG
NCAM_c_2  TAAAGCAGAG AATGGcTTGG GGTCTCCAGC gATAAAGTCA ATCAGAGTGG
NCAM_d_2  TAAAGCAGAG AATGGcTTGG GGTCTCCAGC gATAAAGTCA ATCAGAGTGG
NCAM_d_1  TAAAGCAGAG AATGGcTTGG GGTCTCCAGC gATAAAGTCA ATCAGAGTGG

          351                                                 400
NCAM_c_1  ATGTATACTA TTTGGATGAT CCAGTAGTAA CTGTTCATCA AAGTATAGGT
NCAM_c_2  ATGTATACTA TTTGGATGAT CCAGTAGTAA CTGTTCATCA AAGTATAGGT
NCAM_d_2  ATGTATACTA TTTGGATGAT CCAGTAGTAA CTGTTCATCA AAGTATAGGT
NCAM_d_1  ATGTATACTA TTTGGATGAT CCAGTAGTAA CTGTTCATCA AAGTATAGGT
```

Fig. 1 (Cont.[1])

```
             401                                                  450
NCAM_c_1   GAAGcTAAAg AACAATTTTa CTATGAGAGA ACAGTGTTCC TCCGGTGTGT
NCAM_c_2   GAAGcTAAAg AACAATTTTa CTATGAGAGA ACAGTGTTCC TCCGGTGTGT
NCAM_d_2   GAAGcTAAAg AACAATTTTa CTATGAGAGA ACAGTGTTCC TCCGGTGTGT
NCAM_d_1   GAAGcTAAAg AACAATTTTa CTATGAGAGA ACAGTGTTCC TCCGGTGTGT

             451                                                  500
NCAM_c_1   TGCCAATTCC AATCCTCCTG TTcGGTATAG CTGGAGACGT GGCCAGGAGG
NCAM_c_2   TGCCAATTCC AATCCTCCTG TTcGGTATAG CTGGAGACGT GGCCAGGAGG
NCAM_d_2   TGCCAATTCC AATCCTCCTG TTcGGTATAG CTGGAGACGT GGCCAGGAGG
NCAM_d_1   TGCCAATTCC AATCCTCCTG TTcGGTATAG CTGGAGACGT GGCCAGGAGG

             501                                                  550
NCAM_c_1   TCTTGCTGCA AGGATCTGAT AAAGGAGTTG AGATTTATGA ACCATTCTTT
NCAM_c_2   TCTTGCTGCA AGGATCTGAT AAAGGAGTTG AGATTTATGA ACCATTCTTT
NCAM_d_2   TCTTGCTGCA AGGATCTGAT AAAGGAGTTG AGATTTATGA ACCATTCTTT
NCAM_d_1   TCTTGCTGCA AGGATCTGAT AAAGGAGTTG AGATTTATGA ACCATTCTTT

             551                                                  600
NCAM_c_1   ACCCAGGGTG AAACAAAGAT CTTAAAACTA AAGAATCTTC GACCTCAGGA
NCAM_c_2   ACCCAGGGTG AAACAAAGAT CTTAAAACTA AAGAATCTTC GACCTCAGGA
NCAM_d_2   ACCCAGGGTG AAACAAAGAT CTTAAAACTA AAGAATCTTC GACCTCAGGA
NCAM_d_1   ACCCAGGGTG AAACAAAGAT CTTAAAACTA AAGAATCTTC GACCTCAGGA
```

Fig. 1 (Cont.[2])

```
                                                650
NCAM_c_1  CTATGCTAAT TATAGCTGCA TTGCTTCAGT GAGGAATGTA TGTAATATTC
NCAM_c_2  CTATGCTAAT TATAGCTGCA TTGCTTCAGT GAGGAATGTA TGTAATATTC
NCAM_d_2  CTATGCTAAT TATAGCTGCA TTGCTTCAGT GAGGAATGTA TGTAATATTC
NCAM_d_1  CTATGCTAAT TATAGCTGCA TTGCTTCAGT GAGGAATGTA TGTAATATTC

601                                             700
NCAM_c_1  CTGATAAGAT GGTGTCGTTT AGACTGTCCA ATAAAACAGC ATCACCGTCA
NCAM_c_2  CTGATAAGAT GGTGTCGTTT AGACTGTCCA ATAAAACAGC ATCACCGTCA
NCAM_d_2  CTGATAAGAT GGTGTCGTTT AGACTGTCCA ATAAAACAGC ATCACCGTCA
NCAM_d_1  CTGATAAGAT GGTGTCGTTT AGACTGTCCA ATAAAACAGC ATCACCGTCA

651                                             750
NCAM_c_1  ATTAAACTCT TGGTGGATGA TCCTATAGTT GTAAATCCTG GAGAGGCCAT
NCAM_c_2  ATTAAACTCT TGGTGGATGA TCCTATAGTT GTAAATCCTG GAGAGGCCAT
NCAM_d_2  ATTAAACTCT TGGTGGATGA TCCTATAGTT GTAAATCCTG GAGAGGCCAT
NCAM_d_1  ATTAAACTCT TGGTGGATGA TCCTATAGTT GTAAATCCTG GAGAGGCCAT

701                                             800
NCAM_c_1  AACATTAGTA TGTGTTACAA CAGGAGGAGA GCCTGCACCT TCTCTCACCT
NCAM_c_2  AACATTAGTA TGTGTTACAA CAGGAGGAGA GCCTGCACCT TCTCTCACCT
NCAM_d_2  AACATTAGTA TGTGTTACAA CAGGAGGAGA GCCTGCACCT TCTCTCACCT
NCAM_d_1  AACATTAGTA TGTGTTACAA CAGGAGGAGA GCCTGCACCT TCTCTCACCT
```

Fig. 1 (Cont.[3])

```
          801                                                850
NCAM_c_1  GGGTCAGGTC CTTTGGGACT CTGCCTGAAA AGACTGTTTT GAAtGGAGGA
NCAM_c_2  GGGTCAGGTC CTTTGGGACT CTGCCTGAAA AGACTGTTTT GAAtGGAGGA
NCAM_d_2  GGGTCAGGTC CTTTGGGACT CTGCCTGAAA AGACTGTTTT GAAtGGAGGA
NCAM_d_1  GGGTCAGGTC CTTTGGGACT CTGCCTGAAA AGACTGTTTT GAAtGGAGGA
          851                                                900
NCAM_c_1  ACTTTGACCA TACCTGCCAT CACCTcagat gatgcTGGTA CTTaCAgcTG
NCAM_c_2  ACTTTGACCA TACCTGCCAT CACCTcagat gatgcTGGTA CTTaCAgcTG
NCAM_d_2  ACTTTGACCA TACCTGCCAT CACCTCAGAT GATGCTGGTA CTTACAGCTG
NCAM_d_1  ACTTTGACCA TACCTGCCAT CACCTCAGAT GATGCTGGTA CTTACAGCTG

          901                                                950
NCAM_c_1  CATTGCCAAT AATaATGTGG GaAaCcCTGC AAaAAAGTCC aCcaACATCA
NCAM_c_2  CATTGCCAAT AATaATGTGG GaAaCcCTGC AAaAAAGTCC aCcaACATCA
NCAM_d_2  CATTGCCAAT AATaATGTGG GaAaCCcTGC AAAAAAGTCC ACCAACATCA
NCAM_d_1  CATTGCCAAT AATaATGTGG GaAaCCcTGC AAAAAAGTCC ACCAACATCA

          951                                               1000
NCAM_c_1  TTGtgAGAGC aTtAAAAAAA GgaCGATTTT GGATCACACC AGATCCTTAT
NCAM_c_2  TTGtgAGAGC aTtAAAAAAA GgaCGATTTT GGATCACACC AGATCCTTAT
NCAM_d_2  TTGTGAGAGC ATtAAAAAAA GGaCGATTTT GGATCACACC AgATCCTtAT
NCAM_d_1  TTGTGAGAGC ATtAAAAAAA GGaCGATTTT GGATCACACC AgATCCTtAT
```

Fig. 1(Cont.[4])

```
                                                                1050
NCAM_c_1  CACAAAGATG aCAACATCCA GATTGgCCGT GAGGTGAAAA TATCTTGCCA
NCAM_c_2  CACAAAGATG aCAACATCCA GATTGgCCGT GAGGTGAAAA TATCTTGCCA
NCAM_d_2  CaCAAAGATG ACAACATCCA GATTGGCCGT gAGGTGAaAA tATCTTGCCA
NCAM_d_1  CaCAAAGATG ACAACATCCA GATTGGCCGT gAGGTGAaAA tATCTTGCCA

     1051                                                       1100
NCAM_c_1  AGTAGAAGCT GTTCCTTCTG AGGAGCTAAC ATTTAGTTGG TTTAAAAATG
NCAM_c_2  AGTAGAAGCT GTTCCTTCTG AGGAGCTAAC ATTTAGTTGG TTTAAAAATG
NCAM_d_2  AGTAGAAGCT GTTCCTTCTG AGGAGCTAaC aTTTAGTTGG TTTAAAAATG
NCAM_d_1  AGTAGAAGCT GTTCCTTCTG AGGAGCTAaC aTTTAGTTGG TTTAAAAATG

     1101                                                       1150
NCAM_c_1  GTCGTCCATT AAgAAGTTCT GAGCGGATGG TCATTACACA GACTGATCCT
NCAM_c_2  GTCGTCCATT AAgAAGTTCT GAGCGGATGG TCATTACACA GACTGATCCT
NCAM_d_2  GTCGTCCATT AAgAAGTTCT GAGCGGATGG TCATTACACA GAcTGATCCT
NCAM_d_1  GTCGTCCATT AAgAAGTTCT GAGCGGATGG TCATTACACA GAcTGATCCT

     1151                                                       1200
NCAM_c_1  GATGTCTCTC CGGGAACAaC AAACTTGGAC ATCATTGATT TAAAATTCaC
NCAM_c_2  GATGTCTCTC CGGGAACAaC AAACTTGGAC ATCATTGATT TAAAATTCaC
NCAM_d_2  GATGTCTCTC CGGGAaCAaC AAACTTGGAC ATCATTGATT TAAAATTCaC
NCAM_d_1  GATGTCTCTC CGGGAaCAaC AAACTTGGAC ATCATTGATT TAAAATTCaC
```

Fig. 1(Cont.[5])

```
      1201                                                    1250
NCAM_c_1   ggATTcTGGG acGTaCACAt GtGTAGCATC TCTGAAGGGA GgAGGaATAT
NCAM_c_2   ggATTcTGGG acGTaCACAt GtGTAGCATC TCTGAAGGGA GgAGGaATAT
NCAM_d_2   GGATTcTGGG ACGTACACAT GTGTAGCATC TCTGAAGGGA GGAGGAATAT
NCAM_d_1   GGATTcTGGG ACGTACACAT GTGTAGCATC TCTGAAGGGA GGAGGAATAT

      1251                                                    1300
NCAM_c_1   CTgATATcAG TATCGaTGTT AATAtATCCA GCAGCaCAGT TCCACCCAAT
NCAM_c_2   CTgATATcAG TATCGaTGTT AATAtATCCA GCAGCaCAGT TCCACCCAAT
NCAM_d_2   CTGATATCAG TATCGATGTT AATaTatCCA GCAGCACAGt TCCa......
NCAM_d_1   CTGATATCAG TATCGATGTT AATaTatCCA GCAGCACAGt TCCa......

      1301                                                    1350
NCAM_c_1   CTGACTGTTC CACAGGAAAA ATCACCATTG GTCACCAGAG AAGGAGACAC
NCAM_c_2   CTGACTGTTC CACAGGAAAA ATCACCATTG GTCACCAGAG AAGGAGACAC
NCAM_d_2   .......... .......... .......... .......... ..........
NCAM_d_1   .......... .......... .......... .......... ..........

      1351                                                    1400
NCAM_c_1   AATAGAACTG CAATGTCAAG TAACTGGCAA ACCTAAACCA ATCATCCTTT
NCAM_c_2   AATAGAACTG CAATGTCAAG TAACTGGCAA ACCTAAACCA ATCATCCTTT
NCAM_d_2   .......... .......... .......... .......... ..........
NCAM_d_1   .......... .......... .......... .......... ..........
```

Fig. 1(Cont.[6])

```
                                                                1450
NCAM_c_1  GGTCTAGAGC GGATAAAGAA GTTGCAATGC CTGATGGATC AATGCAAATG
NCAM_c_2  GGTCTAGAGC GGATAAAGAA GTTGCAATGC CTGATGGATC AATGCAAATG
NCAM_d_2  .......... .......... .......... .......... ..........
NCAM_d_1  .......... .......... .......... .......... ..........

1401                                                            1500
NCAM_c_1  gAGAGTTATG ATGGAACACT GAGGATTGTG AATGTATCTA GGGAAATGTC
NCAM_c_2  gAGAGTTATG ATGGAACACT GAGGATTGTG AATGTATCTA GGGAAATGTC
NCAM_d_2  .......... .......... .......... .......... ..........
NCAM_d_1  .......... .......... .......... .......... ..........

1451                                                            1550
NCAM_c_1  AGGAATGTAC AGATGTCAGa CCAGcCAATA CAATGGATTT AACGTGAAAC
NCAM_c_2  AGGAATGTAC AGATGTCAGa CCAGcCAATA CAATGGATTT AACGTGAAAC
NCAM_d_2  .......... .......... .......... .......... ..........
NCAM_d_1  .......... .......... .......... .......... ..........

1501                                                            1600
NCAM_c_1  CAAGGGAAGC CtTGGTGCAG CTCATCGTTC AGTATCCCCC TGCAGTGGAA
NCAM_c_2  CAAGGGAAGC CtTGGTGCAG CTCATCGTTC AGTATCCCCC TGCAGTGGAA
NCAM_d_2  .......... .......... .......... .........CC TGCAGTGGAA
NCAM_d_1  .......... .......... .......... .........CC TGCAGTGGAA
```

Fig. 1(Cont.[7])

```
                                                                1650
NCAM_c_1  CCAGCATTCT TGGAAATCCG GCAAGGACAG GATCGAAGTG TCACTATGAG
NCAM_c_2  CCAGCATTCT TGGAAATCCG GCAAGGACAG GATCGAAGTG TCACTATGAG
NCAM_d_2  CCAGCATTCT TGGAAATCCG GCAAGGACAG GATCGAAGTG TCACTATGAG
NCAM_d_1  CCAGCATTCT TGGAAATCCG GCAAGGACAG GATCGAAGTG TCACTATGAG
1601

                                                                1700
NCAM_c_1  TTGCAGAGTA CTGAGAGCCT ATCCAATACG GGTGCTGACC TATGAGTGGC
NCAM_c_2  TTGCAGAGTA CTGAGAGCCT ATCCAATACG GGTGCTGACC TATGAGTGGC
NCAM_d_2  TTGCAGAGTA CTGAGAGCCT ATCCAATACG GGTGCTGACC TATGAGTGGC
NCAM_d_1  TTGCAGAGTA CTGAGAGCCT ATCCAATACG GGTGCTGACC TATGAGTGGC
1651

                                                                1750
NCAM_c_1  GCTTGGGCAA TAAATTATTA CGGACGGGTC AATTTGACTC TCAGGAATAC
NCAM_c_2  GCTTGGGCAA TAAATTATTA CGGACGGGTC AATTTGACTC TCAGGAATAC
NCAM_d_2  GCTTGGGCAA TAAATTATTA CGGACGGGTC AAtTTGACTC TCAGGAATAC
NCAM_d_1  GCTTGGGCAA TAAATTATTA CGGACGGGTC AAtTTGACTC TCAGGAATAC
1701

                                                                1800
NCAM_c_1  ACAGAGTACG CTGTGAAGAG TCTTTCCAAT GAAAACTATG GGGTTTATAA
NCAM_c_2  ACAGAGTACG CTGTGAAGAG TCTTTCCAAT GAAAACTATG GGGTTTATAA
NCAM_d_2  ACAGAGTACG CTGTGAAGAG TCTTTCCAAT GAAAACTATG GGGTTTATAA
NCAM_d_1  ACAGAGTACG CTGTGAAGAG TCTTTCCAAT GAAAACTATG GGGTTTATAA
1751
```

Fig. 1(Cont.[8])

```
                                                                1850
NCAM_c_1  CTGTAGCATC ATAAATGAAG CTGGAGCTGG GAGATGCAGC TTTCTTGTTA
NCAM_c_2  CTGTAGCATC ATAAATGAAG CTGGAGCTGG GAGATGCAGC TTTCTTGTTA
NCAM_d_2  CTGTAGCATC ATAAATGAAG CTGGAGCTGG GAGATGCAGC TTTCTTGTTA
NCAM_d_1  CTGTAGCATC ATAAATGAAG CTGGAGCTGG GAGATGCAGC TTTCTTGTTA

                                                                1900
NCAM_c_1  CAGGAAAGGC CTATGCTCCA GAATTCTATT ATGATACCTA CAATCCAGTA
NCAM_c_2  CAGGAAAGGC CTATGCTCCA GAATTCTATT ATGATACCTA CAATCCAGTA
NCAM_d_2  CAGGAAAGGC CTATGCTCCA GAATTCTATT ATGATACCTA CAATCCAGTA
NCAM_d_1  CAGGAAAGGC CTATGCTCCA GAATTCTATT ATGATACCTA CAATCCAGTA

                                                                1950
NCAM_c_1  TGGCAGAACA GACACCGTGT TTATTCTTAC AGTCTACAGT GGACACAGAT
NCAM_c_2  TGGCAGAACA GACACCGTGT TTATTCTTAC AGTCTACAGT GGACACAGAT
NCAM_d_2  TGGCAGAACA GACACCGTGT TTATTCTTAC AGTCTACAGT GGACACAGAT
NCAM_d_1  TGGCAGAACA GACACCGTGT TTATTCTTAC AGTCTACAGT GGACACAGAT

                                                                2000
NCAM_c_1  GAATCCTGAT GCAGTGGATC GGATTGTTGC ATACCGGTTG GGCATCAGGC
NCAM_c_2  GAATCCTGAT GCAGTGGATC GGATTGTTGC ATACCGGTTG GGCATCAGGC
NCAM_d_2  GAATCCTGAT GCAGTGGATC GGATTGTTGC ATACCGGTTG GGCATCAGGC
NCAM_d_1  GAATCCTGAT GCAGTGGATC GGATTGTTGC ATACCGGTTG GGCATCAGGC
```

Fig. 1(Cont.9)

```
     2001                                                2050
NCAM_c_1  AGGCTGGACA GCAGCGCTGG TGGGAGCAGG AGATTAAAAT AAATGGGAAT
NCAM_c_2  AGGCTGGACA GCAGCGCTGG TGGGAGCAGG AGATTAAAAT AAATGGGAAT
NCAM_d_2  AGGCTGGACA GCAGCGCTGG TGGGAGCAGG AGATTAAAAT AtATGGGAAT
NCAM_d_1  AGGCTGGACA GCAGCGCTGG TGGGAGCAGG AGATTAAAAT AtATGGGAAT

     2051                                                2100
NCAM_c_1  ATTCAAAAGG GAGAATTAAT TACATATAAC TTGACAGAGC TAATTAAACC
NCAM_c_2  ATTCAAAAGG GAGAATTAAT TACATATAAC TTGACAGAGC TAATTAAACC
NCAM_d_2  ATTCAAAAGG GAGAATTAAT TACATATAAC TTGACAGAGC TAATTAAACC
NCAM_d_1  ATTCAAAAGG GAGAATTAAT TACATATAAC TTGACAGAGC TAATTAAACC

     2101                                                2150
NCAM_c_1  AGAAGCTTAT GAAGTCCGAC TGACTCCTCT CACCAAATTT GGTGAAGGAG
NCAM_c_2  AGAAGCTTAT GAAGTCCGAC TGACTCCTCT CACCAAATTT GGTGAAGGAG
NCAM_d_2  AGAAGCTTAT GAAGTCCGAC TGACTCCTCT CACCAAATTT GGTGAAGGAG
NCAM_d_1  AGAAGCTTAT GAAGTCCGAC TGACTCCTCT CACCAAATTT GGTGAAGGAG
     2151                                                2200
NCAM_c_1  ATTCAACAAT TCGTGTGATC AAATATAGTG CTCCTGTAAA TCCTCATTTG
NCAM_c_2  ATTCAACAAT TCGTGTGATC AAATATAGTG CTCCTGTAAA TCCTCATTTG
NCAM_d_2  ATTCAACAAT TCGTGTGATC AAATATAGTG CTCCTGTAAA TCCTCATTTG
NCAM_d_1  ATTCAACAAT TCGTGTGATC AAATATAGTG CTCCTGTAAA TCCTCATTTG
```

Fig. 1 (Cont.[10])

```
            2201                                                2250
NCAM_c_1   AGAGAATTTC ATTGTGGATT TGAAGATGGT AATATTTGTT TGTTCACTCA
NCAM_c_2   AGAGAATTTC ATTGTGGATT TGAAGATGGT AATATTTGTT TGTTCACTCA
NCAM_d_2   AGAGAATTTC ATTGTGGATT TGAAGATGGT AATATTTGTT TGTTCACTCA
NCAM_d_1   AGAGAATTTC ATTGTGGATT TGAAGATGGT AATATTTGTT TGTTCACTCA

            2251                                                2300
NCAM_c_1   AGATGATACA GATAATTTTG ACTGGACAAA GCAAAGTACA GCAACAAGAA
NCAM_c_2   AGATGATACA GATAATTTTG ACTGGACAAA GCAAAGTACA GCAACAAGAA
NCAM_d_2   AGATGATACA GATAATTTTG ACTGGACAAA GCAAAGTACA GCAACAAGAA
NCAM_d_1   AGATGATACA GATAATTTTG ACTGGACAAA GCAAAGTACA GCAACAAGAA

            2301                                                2350
NCAM_c_1   ATACAAAATA TACTCCTAAT ACAGGACCTA ATGCTGACCG TAGTGGCTCC
NCAM_c_2   ATACAAAATA TACTCCTAAT ACAGGACCTA ATGCTGACCG TAGTGGCTCC
NCAM_d_2   ATACAAAATA TACTCCTAAT ACAGGACCTA ATGCTGACCG TAGTGGCTCC
NCAM_d_1   ATACAAAATA TACTCCTAAT ACAGGACCTA ATGCTGACCG TAGTGGCTCC

            2351                                                2400
NCAM_c_1   AAAGAAGGTT TTTATATGTA CATTGAGACA TCACGACCCA GATTGGaAGG
NCAM_c_2   AAAGAAGGTT TTTATATGTA CATTGAGACA TCACGACCCA GATTGGaAGG
NCAM_d_2   AAAGAAGGTT TTTATATGTA CATTGAGACA TCACGACCCA GATTGGaAGG
NCAM_d_1   AAAGAAGGTT TTTATATGTA CATTGAGACA TCACGACCCA GATTGGaAGG
```

Fig. 1 (Cont.[11])

```
                                                                2450
2401
NCAM_c_1  CGAAAAGGCT CGACTTCTCA GCCCTGTTTT CAGCATAGCT CCCaAAAACC
NCAM_c_2  CGAAAAGGCT CGACTTCTCA GCCCTGTTTT CAGCATAGCT CCCaAAAACC
NCAM_d_2  CGAAAAGGCT CGACTTCTCA GCCCTGTTTT CAGCATAGCT CCCaAAAACC
NCAM_d_1  CGAAAAGGCT CGACTTCTCA GCCCTGTTTT CAGCATAGCT CCCaAAAACC

                                                                2500
2451
NCAM_c_1  CTTATGGACC CaCAAACACT GCATATTGTT TCAGCTTCTT TTATCACATG
NCAM_c_2  CTTATGGACC CaCAAACACT GCATATTGTT TCAGCTTCTT TTATCACATG
NCAM_d_2  CTTATGGACC CaCAAACACT GCATATTGTT TCAGCTTCTT TTATCACATG
NCAM_d_1  CTTATGGACC CaCAAACACT GCATATTGTT TCAGCTTCTT TTATCACATG

                                                                2550
2501
NCAM_c_1  TATGGACAAC ATATAGGTGT CTTAAATGTT TATCTACGTT TGAAAGGGCA
NCAM_c_2  TATGGACAAC ATATAGGTGT CTTAAATGTT TATCTACGTT TGAAAGGGCA
NCAM_d_2  TATGGACAAC ATATAGGTGT CTTAAATGTT TATCTACGTT TGAAAGGGCA
NCAM_d_1  TATGGACAAC ATATAGGTGT CTTAAATGTT TATCTACGTT TGAAAGGGCA

                                                                2600
2551
NCAM_c_1  AACAACAATA GAGAATCCAC TGTGGTCTTC AAGTGGGAAT AAAGGACAAA
NCAM_c_2  AACAACAATA GAGAATCCAC TGTGGTCTTC AAGTGGGAAT AAAGGACAAA
NCAM_d_2  AACAACAATA GAGAATCCAC TGTGGTCTTC AAGTGGGAAT AAAGGACAAA
NCAM_d_1  AACAACAATA GAGAATCCAC TGTGGTCTTC AAGTGGGAAT AAAGGACAAA
```

Fig. 1 (Cont.[12])

```
                2650
2601
NCAM_c_1  GATGGAATGA GGCTCATGTT AATATATACC CAATTACTTC ATTTCAGCTC
NCAM_c_2  GATGGAATGA GGCTCATGTT AATATATACC CAATTACTTC ATTTCAGCTC
NCAM_d_2  GATGGAATGA GGCTCATGTT AATATATACC CAATTACTTC ATTTCAGCTC
NCAM_d_1  GATGGAATGA GGCTCATGTT AATATATACC CAATTACTTC ATTTCAGCTC

                2700
2651
NCAM_c_1  ATTTTtGAAG GTATCCGAGG TCCTGGAATA GAAGGTGAcA TTGCTATTGA
NCAM_c_2  ATTTTtGAAG GTATCCGAGG TCCTGGAATA GAAGGTGAcA TTGCTATTGA
NCAM_d_2  ATTTTtGAAG GTATCCGAGG TCCTGGAATA GAAGGTGAcA TTGCTATTGA
NCAM_d_1  ATTTTtGAAG GTATCCGAGG TCCTGGAATA GAAGGTGAcA TTGCTATTGA

                2750
2701
NCAM_c_1  TGATGTATCA ATTGCAGAAG GAGAATGTGC AAAACAAGAC CTAGCAACTA
NCAM_c_2  TGATGTATCA ATTGCAGAAG GAGAATGTGC AAAACAAGAC CTAGCAACTA
NCAM_d_2  TGATGTATCA ATTGCAGAAG GAGAATGTGC AAAACAAGAC CTAGCAACTA
NCAM_d_1  TGATGTATCA ATTGCAGAAG GAGAATGTGC AAAACAAGAC CTAGCAACTA

                2800
2751
NCAM_c_1  AGAATTCCGT TGATGGTGCT GTTGGGATTT TGGTTCATAT ATGGCTTTTT
NCAM_c_2  AGAATTCCGT TGATGGTGCT GTTGGGATTT TGGTTCATAT ATGGCTTTTT
NCAM_d_2  AGAATTCCGT TGATGGTGCT GTTGGGATTT TGGTTCATAT ATGGCTTTTT
NCAM_d_1  AGAATTCCGT TGATGGTGCT GTTGGGATTT TGGTTCATAT ATGGCTTTTT
```

Fig. 1 (Cont.[13])

```
                                                                2850
NCAM_c_1  CCCATTATCG TCCTCATCTC TATCTTAAGT CCTCGAAGGT GACCTTATCC
NCAM_c_2  CCCATTATCG TCCTCATCTC TATCTTAAGT CCTCGAAGGT GACCTTATCC
NCAM_d_2  CCCATTATCG TCCTCATCTC TATCTTAAGT CCTCGAAGGT GACCTTATCC
NCAM_d_1  CCCATTATCG TCCTCATCTC TATCTTAAGT CCTCGAAGGT GACCTTATCC
     2801
     2851                                                       2900
NCAM_c_1  TGGCAGAGGC TATAAAAGAT TCACCAGGCA CTGGCATGAA GAAAGAGTCT
NCAM_c_2  TGGCAGAGGC TATAAAAGAT TCACCAGGCA CTGGCATGAA GAAAGAGTCT
NCAM_d_2  TGGCAGAGGC TATAAAAGAT TCACCAGGCA CTGGCATGAA GAAAGAGTCT
NCAM_d_1  TGGCAGAGGC TATAAAAGAT TCACCAGGCA CTGGCATGAA GAAAGAGTCT
     2901                                                       2950
NCAM_c_1  TTGTAAATGG ACATTGAACA AACAAACTAC CAAAGATTCC TCCACTGACT
NCAM_c_2  TTGTAAATGG ACATTGAACA AACAAACTAC CAAAGATTCC TCCACTGACT
NCAM_d_2  TTGTAAATGG ACATTGAACA AACAAACTAC CAAAGATTCC TCCACTGACT
NCAM_d_1  TTGTAAATGG ACATTGAACA AACAAACTAC CAAAGATTCC TCCACTGACT
     2951                                                       3000
NCAM_c_1  ACNGGACTCA AAAATAAAAT AATAAAAACA AATTTTTTTA AGCACCTGGG
NCAM_c_2  ACNGGACTCA AAAATAAAAT AATAAAAACA AATTTTTTTA AGCACCTGGG
NCAM_d_2  ACNGGACTCA AAAATAAAAT AATAAAAACA AATTTTTTTA AGCACCTGGG
NCAM_d_1  ACNGGACTCA AAAATAAAAT AATAAAAACA AATTTTTTTA AGCACCTGGG
```

Fig. 1 (Cont.[14])

```
                3001                                                3050
NCAM_c_1   GATAAAAAGA CATCATGGAA GTATAACTTA TTNCAGACTA AACATAAANG
NCAM_c_2   GATAAAAAGA CATCATGGAA GTATAACTTA TTNCAGACTA AACATAAANG
NCAM_d_2   GATAAAAAGA CATCATGGAA GTATAACTTA TTNCAGACTA AACATAAANG
NCAM_d_1   GATAAAAAGA CATCATGGAA GTATAACTTA TTNCAGACTA AACATAAANG

                3051                3065
NCAM_c_1   ATAATCGTTG ACCTG
NCAM_c_2   ATAATCGTTG ACCTG
NCAM_d_2   ATAATCGTTG ACCTG
NCAM_d_1   ATAATCGTTG ACCTG
```

Fig. 1 (Cont.[15])

```
   6 TTTCAGCATAGCTCCCAAAAATCCATATGGACCTACAAATAGTGCATATT  55
     ||||||||||||||||||||||||| || |||||||| |||||| | ||||||||||
2423 TTTCAGCATAGCTCCCaAAAACCCTTATGGACCCaCAAACACTGCATATT 2472

  56 GTTTCAGTTTCTTCTATCACATGTACGGGCAACATATAGGGGTTTTAAAT 105
     |||||||| |||||| |||||||||||| || |||||||||||| || ||||||
2473 GTTTCAGCTTCTTTTATCACATGTATGGACAACATATAGGTGTCTTAAAT 2522

 106 GTATATCTACGTTTGAAAGGGCAGACAACGATAGAGAATCCGCTATGGTC 155
     || |||||||||||||||||||| |||||| |||||||||||| || ||||||
2523 GTTTATCTACGTTTGAAAGGGCAAACAACAATAGAGAATCCACTGTGGTC 2572

 156 GTCGAGTGGGAACAAAGGACAACGATGGAATGAAGCTCATGTTAATATAT 205
     || |||||||| |||||||||| |||||||||| |||||||||||||||||
2573 TTCAAGTGGGAATAAAGGACAAAGATGGAATGAGGCTCATGTTAATATAT 2622

 206 ATCCAATT.CTTCATTTCAGTTAATTTTTGAAGGCATTCGAGGTCCTGGG 254
     | |||||||| |||||||||||| | |||||||||||| || ||||||||||||
2623 ACCCAATTACTTCATTTCAGCTCATTTTTtGAAGGTATCCGAGGTCCTGGA 2672
```

Fig. 2

```
255  ATAGAGGGTGACATCGCCATTGATGATGTATCAATTGCTGAAGGAGAATG  304
     |||||| |||||||||| || |||||||||||||||||| ||||||||||
2673 ATAGAAGGTGACATTGCTATTGATGATGTATCAATTGCAGAAGGAGAATG  2722

305  TGCAAAACAAGACCTACCAACTAAGAATTCCGTGGATGGTGCTGTTGGGA  354
     |||||||||||||||||||||||||||||| |||||||||||||||||||
2723 TGCAAAACAAGACCTAGCAACTAAGAATTCCGTTGATGGTGCTGTTGGGA  2772

355  TCTTTAGTTCATATATGGCTTTTTTCCAGTTATCATCCTCATCTCTATCTTA  404
     | || |||||||||||||||||||  ||||| |||||||||||||||||||
2773 TTTTGGTTCATATATGGCTTTTTTCCCATTATCGTCCTCATCTCTATCTTA  2822

405  AGCCCTCGAAGGTGACCTTATCCTGGCAGAGGCTATTAATGATTCACCAG  454
     || |||||||||||||||||||||||||||||| || ||||||||||||
2823 AGTCCTCGAAGGTGACCTTATCCTGGCAGAGGCTATAAAGATTCACCAG  2872

455  GCACTGGCATGAAGAAGAAGCTTTGTAAATGGACACTGAAAAACAAAC  504
     |||||||||||||||||||  |||||||||||||||||| ||||||||||
2873 GCACTGGCATGAAGAAAGAGTCTTTGTAAATGGACATTGAACAAACAAAC  2922
```

Fig. 2 (Cont.[1])

```
505  TACCAAAGATTCATCCACTGACTACT.GACTCAAAAATTAAATAATAAAA 553
     ||||||||||||||||||||||||:|||||||||||||||||||||||||
2923 TACCAAAGATTCCTCCACTGACTACNGGACTCAAAAATAAAATAATAAAA 2972

554  ACAACATTTTTTTAAGCA.CTGGGGATTAAAAAAAAAGACATCATGGAAA 602
     |||||||||||||||||| |||||||||||||||||||||||||||||
2973 ACAA.ATTTTTTTAAGCACCTGGGGGAT.....AAAAAGACATCATGGAAG 3016

603  TATAACTTATTCCAAACTACATATAGAAGATATTC.TTGGCCTG 645
     ||||||||||||:||||||||||:|||||||||||||||||||
3017 TATAACTTATTNCAGACTAAACATAAANGATAATCGTTGACCTG 3060
```

Fig. 2 (Cont.[2])

```
  1 FFTNKAEDQRCLASAPPTVRIVHSGLACNIEEERYSERVYTIREGETLEL 50
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  1 FFTNKAEDQRCLASAPPTVRIVHSGLACNIEEERYSERVYTIREGETLEL 50

 51 TCLVTGHPRPQIRWTKTAGSASDRFQDSSVFNETLRITNIQRHQGGRYYC 100
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 51 TCLVTGHPRPQIRWTKTAGSASDRFQDSSVFNETLRITNIQRHQGGRYYC 100

101 KAENGLGSPAIKSIRVDVYYLDDPVVTVHQSIGEAKEQFYYERTVFLRCV 150
    |||||||||||||||||||||||||||||||||||||||||||||||||||
101 KAENGLGSPAIKSIRVDVYYLDDPVVTVHQSIGEAKEQFYYERTVFLRCV 150

151 ANSNPPVRYSWRRGQEVLLQGSDKGVEIYEPFFTQGETKILKLKNLRPQD 200
    |||||||||||||||||||||||||||||||||||||||||||||||||||
151 ANSNPPVRYSWRRGQEVLLQGSDKGVEIYEPFFTQGETKILKLKNLRPQD 200

201 YANYSCIASVRNVCNIPDKMVSFRLSNKTASPSIKLLVDDPIVVNPGEAI 250
    |||||||||||||||||||||||||||||||||||||||||||||||||||
201 YANYSCIASVRNVCNIPDKMVSFRLSNKTASPSIKLLVDDPIVVNPGEAI 250
```

Fig. 3

```
251 TLVCVTTGGEPAPSLTWVRSFGTLPEKTVLNGGTLTIPAITSDDAGTYSC 300
    |||||||||||||||||||||||||||||||||||||||||||||||||||
251 TLVCVTTGGEPAPSLTWVRSFGTLPEKTVLNGGTLTIPAITSDDAGTYSC 300

301 IANNNVGNPAKKSTNIIVRALKKGRFWITPDPYHKDDNIQIGREVKISCQ 350
    |||||||||||||||||||||||||||||||||||||||||||||||||||
301 IANNNVGNPAKKSTNIIVRALKKGRFWITPDPYHKDDNIQIGREVKISCQ 350

351 VEAVPSEELTFSWFKNGRPLRSSERMVITQTDPDVSPGTTNLDIIDLKFT 400
    |||||||||||||||||||||||||||||||||||||||||||||||||||
351 VEAVPSEELTFSWFKNGRPLRSSERMVITQTDPDVSPGTTNLDIIDLKFT 400

401 DSGTYTCVASLKGGGISDISIDVNISSSTV.....................  430
    |||||||||||||||||||||||||||||||
401 DSGTYTCVASLKGGGISDISIDVNISSSTVPPNLTVPQEKSPLVTREGDT 450

431 .....................PPAVEPAFLEIRQGQDRSVTMS 452
                          |||||||||||||||||||||||
501 GMYRCQTSQYNGFNVKPREALVQLIVQYPPAVEPAFLEIRQGQDRSVTMS 550
```

Fig. 3(Cont.[1])

```
453 CRVLRAYPIRVLTYEWRLGNKLLRTGQFDSQEYTEYAVKSLSNENYGVYN 502
    |||||||||||||||||||||||||||||||||||||||||||||||||||
551 CRVLRAYPITRVLTYEWRLGNKLLRTGQFDSQEYTEYAVKSLSNENYGVYN 600

503 CSIINEAGAGRCSFLVTGKAYAPEFYYDTYNPVWQNRHRVYSYSLQWTQM 552
    |||||||||||||||||||||||||||||||||||||||||||||||||||
601 CSIINEAGAGRCSFLVTGKAYAPEFYYDTYNPVWQNRHRVYSYSLQWTQM 650

553 NPDAVDRIVAYRLGIRQAGQQRWWEQEIKINGNIQKGELITYNLTELIKP 602
    |||||||||||||||||||||||||||||||||||||||||||||||||||
651 NPDAVDRIVAYRLGIRQAGQQRWWEQEIKINGNIQKGELITYNLTELIKP 700

603 EAYEVRLTPLTKFGEGDSTIRVIKYSAPVNPHLREFHCGFEDGNICLFTQ 652
    |||||||||||||||||||||||||||||||||||||||||||||||||||
701 EAYEVRLTPLTKFGEGDSTIRVIKYSAPVNPHLREFHCGFEDGNICLFTQ 750

653 DDTDNFDWTKQSTATRNTKYTPNTGPNADRSGSKEGFYMYIETSRPRLEG 702
    |||||||||||||||||||||||||||||||||||||||||||||||||||
751 DDTDNFDWTKQSTATRNTKYTPNTGPNADRSGSKEGFYMYIETSRPRLEG 800
```

Fig. 3(Cont.2)

```
703 EKARLLSPVFSIAPKNPYGPTNTAYCFSFFYHMYGQHIGVLNVYLRLKGQ 752
    |||||||||||||||||||||||||||||||||||||||||||||||||||
801 EKARLLSPVFSIAPKNPYGPTNTAYCFSFFYHMYGQHIGVLNVYLRLKGQ 850

753 TTIENPLWSSSGNKGQRWNEAHVNIYPITSFQLIFEGIRGPGIEGDIAID 802
    |||||||||||||||||||||||||||||||||||||||||||||||||||
851 TTIENPLWSSSGNKGQRWNEAHVNIYPITSFQLIFEGIRGPGIEGDIAID 900

803 DVSIAEGECAKQDLATKNSVDGAVGILVHIWLFPIIVLISILSPRR*PYP 852
    |||||||||||||||||||||||||||||||||||||||||||||||||||
901 DVSIAEGECAKQDLATKNSVDGAVGILVHIWLFPIIVLISILSPRR*PYP 950

853 GRGYKRFTRHWHEERVFVNGH*TNKLPKIPPLTTGLKNKIIKTNFFKHLG 902
    |||||||||||||||||||||||||||||||||||||||||||||||||||
951 GRGYKRFTRHWHEERVFVNGH*TNKLPKIPPLTTGLKNKIIKTNFFKHLG 1000

903 IKRHHGSITYXRLNIXDNR*P 923
    |||||||||||||||||||||
1001 IKRHHGSITYXRLNIXDNR*P 1021
```

Fig. 3(Cont.[3])

```
NCAM_d_1  FFTNKAEDQR CLASAPPTVR IVHSGLACNI EEERYSERVY TIREGETLEL
NCAM_d_2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NCAM_c_1  FFTNKAEDQR CLASAPPTVR IVHSGLACNI EEERYSERVY TIREGETLEL
NCAM_c_2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

       51                                          100
NCAM_d_1  TCLVTGHPRP QIRWTKTAGS ASDRFQDSSV FNETLRITNI QRHQGGRYYC
NCAM_d_2  ~~~~~SQDIH VHRWTKTAGS ASDRFQDSSV FNETLRITNI QRHQGGRYYC
NCAM_c_1  TCLVTGHPRP QIRWTKTAGS ASDRFQDSSV FNETLRITNI QRHQGGRYYC
NCAM_c_2  ~~~~~SQDIH VHRWTKTAGS ASDRFQDSSV FNETLRITNI QRHQGGRYYC

      101                                          150
NCAM_d_1  KAENGLGSPA IKSIRVDVYY LDDPVVTVHQ SIGEAKEQFY YERTVFLRCV
NCAM_d_2  KAENGLGSPA IKSIRVDVYY LDDPVVTVHQ SIGEAKEQFY YERTVFLRCV
NCAM_c_1  KAENGLGSPA IKSIRVDVYY LDDPVVTVHQ SIGEAKEQFY YERTVFLRCV
NCAM_c_2  KAENGLGSPA IKSIRVDVYY LDDPVVTVHQ SIGEAKEQFY YERTVFLRCV

      151                                          200
NCAM_d_1  ANSNPPVRYS WRRGQEVLLQ GSDKGVEIYE PFFTQGETKI LKLKNLRPQD
NCAM_d_2  ANSNPPVRYS WRRGQEVLLQ GSDKGVEIYE PFFTQGETKI LKLKNLRPQD
NCAM_c_1  ANSNPPVRYS WRRGQEVLLQ GSDKGVEIYE PFFTQGETKI LKLKNLRPQD
NCAM_c_2  ANSNPPVRYS WRRGQEVLLQ GSDKGVEIYE PFFTQGETKI LKLKNLRPQD
```

Fig. 4

```
                                                                250
NCAM_d_1  YANYSCIASV RNVCNIPDKM VSFRLSNKTA SPSIKLLVDD PIVVNPGEAI
NCAM_d_2  YANYSCIASV RNVCNIPDKM VSFRLSNKTA SPSIKLLVDD PIVVNPGEAI
NCAM_c_1  YANYSCIASV RNVCNIPDKM VSFRLSNKTA SPSIKLLVDD PIVVNPGEAI
NCAM_c_2  YANYSCIASV RNVCNIPDKM VSFRLSNKTA SPSIKLLVDD PIVVNPGEAI
      201
                                                                300
NCAM_d_1  TLVCVTTGGE PAPSLTWVRS FGTLPEKTVL NGGTLTIPAI TSDDAGTYSC
NCAM_d_2  TLVCVTTGGE PAPSLTWVRS FGTLPEKTVL NGGTLTIPAI TSDDAGTYSC
NCAM_c_1  TLVCVTTGGE PAPSLTWVRS FGTLPEKTVL NGGTLTIPAI TSDDAGTYSC
NCAM_c_2  TLVCVTTGGE PAPSLTWVRS FGTLPEKTVL NGGTLTIPAI TSDDAGTYSC
      251
                                                                350
NCAM_d_1  IANNNVGNPA KKSTNIIVRA LKKGRFWITP DPYHKDDNIQ IGREVKISCQ
NCAM_d_2  IANNNVGNPA KKSTNIIVRA LKKGRFWITP DPYHKDDNIQ IGREVKISCQ
NCAM_c_1  IANNNVGNPA KKSTNIIVRA LKKGRFWITP DPYHKDDNIQ IGREVKISCQ
NCAM_c_2  IANNNVGNPA KKSTNIIVRA LKKGRFWITP DPYHKDDNIQ IGREVKISCQ
      301
                                                                400
NCAM_d_1  VEAVPSEELT FSWFKNGRPL RSSERMVITQ TDPDVSPGTT NLDIIDLKFT
NCAM_d_2  VEAVPSEELT FSWFKNGRPL RSSERMVITQ TDPDVSPGTT NLDIIDLKFT
NCAM_c_1  VEAVPSEELT FSWFKNGRPL RSSERMVITQ TDPDVSPGTT NLDIIDLKFT
NCAM_c_2  VEAVPSEELT FSWFKNGRPL RSSERMVITQ TDPDVSPGTT NLDIIDLKFT
      351
```

Fig. 4 (Cont.[1])

```
         401                                                      450
NCAM_d_1  DSGTYTCVAS LKGGGISDIS IDVNISSSTV .......... ..........
NCAM_d_2  DSGTYTCVAS LKGGGISDIS IDVNISSSTV .......... ..........
NCAM_c_1  DSGTYTCVAS LKGGGISDIS IDVNISSSTV PPNLTVPQEK SPLVTREGDT
NCAM_c_2  DSGTYTCVAS LKGGGISDIS IDVNISSSTV PPNLTVPQEK SPLVTREGDT

         451                                                      500
NCAM_d_1  .......... .......... .......... .......... ..........
NCAM_d_2  .......... .......... .......... .......... ..........
NCAM_c_1  IELQCQVTGK PKPIILWSRA DKEVAMPDGS MQMESYDGTL RIVNVSREMS
NCAM_c_2  IELQCQVTGK PKPIILWSRA DKEVAMPDGS MQMESYDGTL RIVNVSREMS

         501                                                      550
NCAM_d_1  .......... .......... ........PP AVEPAFLEIR QGQDRSVTMS
NCAM_d_2  .......... .......... ........PP AVEPAFLEIR QGQDRSVTMS
NCAM_c_1  GMYRCQTSQY NGFNVKPREA LVQLIVQYPP AVEPAFLEIR QGQDRSVTMS
NCAM_c_2  GMYRCQTSQY NGFNVKPREA LVQLIVQYPP AVEPAFLEIR QGQDRSVTMS

         551                                                      600
NCAM_d_1  CRVLRAYPIR VLTYEWRLGN KLLRTGQFDS QEYTEYAVKS LSNENYGVYN
NCAM_d_2  CRVLRAYPIR VLTYEWRLGN KLLRTGQFDS QEYTEYAVKS LSNENYGVYN
NCAM_c_1  CRVLRAYPIR VLTYEWRLGN KLLRTGQFDS QEYTEYAVKS LSNENYGVYN
NCAM_c_2  CRVLRAYPIR VLTYEWRLGN KLLRTGQFDS QEYTEYAVKS LSNENYGVYN
```

Fig. 4 (Cont.[2])

```
                                                650
NCAM_d_1  CSIINEAGAG RCSFLVTGKA YAPEFYYDTY NPVWQNRHRV YSYSLQWTQM
NCAM_d_2  CSIINEAGAG RCSFLVTGKA YAPEFYYDTY NPVWQNRHRV YSYSLQWTQM
NCAM_c_1  CSIINEAGAG RCSFLVTGKA YAPEFYYDTY NPVWQNRHRV YSYSLQWTQM
NCAM_c_2  CSIINEAGAG RCSFLVTGKA YAPEFYYDTY NPVWQNRHRV YSYSLQWTQM

          651                                   700
NCAM_d_1  NPDAVDRIVA YRLGIRQAGQ QRWWEQEIKI YGNIQKGELI TYNLTELIKP
NCAM_d_2  NPDAVDRIVA YRLGIRQAGQ QRWWEQEIKI YGNIQKGELI TYNLTELIKP
NCAM_c_1  NPDAVDRIVA YRLGIRQAGQ QRWWEQEIKI NGNIQKGELI TYNLTELIKP
NCAM_c_2  NPDAVDRIVA YRLGIRQAGQ QRWWEQEIKI NGNIQKGELI TYNLTELIKP

          701                                   750
NCAM_d_1  EAYEVRLTPL TKFGEGDSTI RVIKYSAPVN PHLREFHCGF EDGNICLFTQ
NCAM_d_2  EAYEVRLTPL TKFGEGDSTI RVIKYSAPVN PHLREFHCGF EDGNICLFTQ
NCAM_c_1  EAYEVRLTPL TKFGEGDSTI RVIKYSAPVN PHLREFHCGF EDGNICLFTQ
NCAM_c_2  EAYEVRLTPL TKFGEGDSTI RVIKYSAPVN PHLREFHCGF EDGNICLFTQ

          751                                   800
NCAM_d_1  DDTDNFDWTK QSTATRNTKY TPNTGPNADR SGSKEGFYMY IETSRPRLEG
NCAM_d_2  DDTDNFDWTK QSTATRNTKY TPNTGPNADR SGSKEGFYMY IETSRPRLEG
NCAM_c_1  DDTDNFDWTK QSTATRNTKY TPNTGPNADR SGSKEGFYMY IETSRPRLEG
NCAM_c_2  DDTDNFDWTK QSTATRNTKY TPNTGPNADR SGSKEGFYMY IETSRPRLEG
```

Fig. 4 (Cont.^{3})

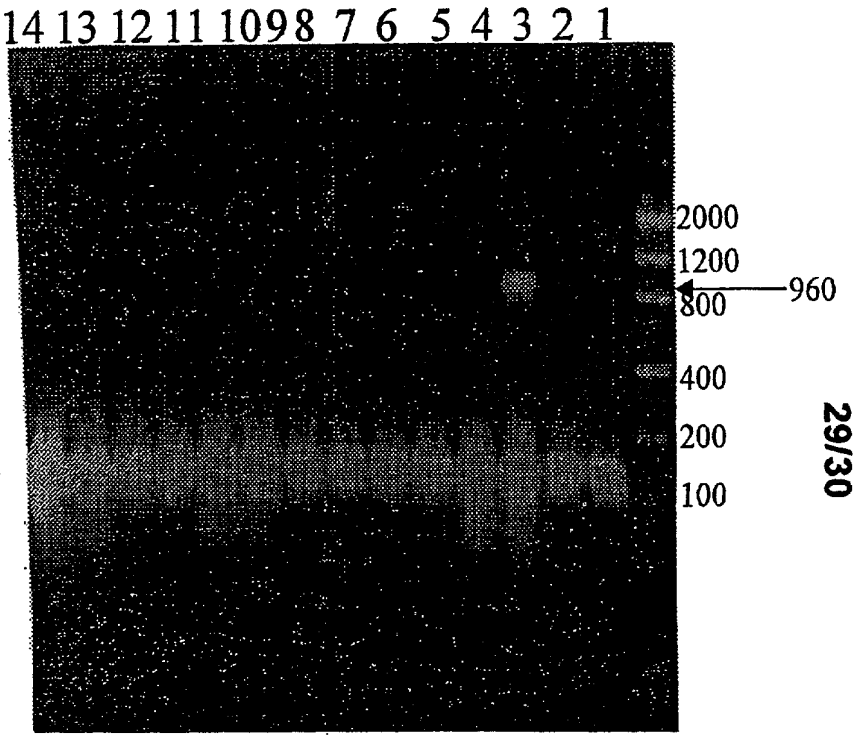
```
                                                                850
NCAM_d_1  EKARLLSPVF SIAPKNPYGP TNTAYCFSFF YHMYGQHIGV LNVYLRLKGQ
NCAM_d_2  EKARLLSPVF SIAPKNPYGP TNTAYCFSFF YHMYGQHIGV LNVYLRLKGQ
NCAM_c_1  EKARLLSPVF SIAPKNPYGP TNTAYCFSFF YHMYGQHIGV LNVYLRLKGQ
NCAM_c_2  EKARLLSPVF SIAPKNPYGP TNTAYCFSFF YHMYGQHIGV LNVYLRLKGQ

      851                                                       900
NCAM_d_1  TTIENPLWSS SGNKGQRWNE AHVNIYPITS FQLIFEGIRG PGIEGDIAID
NCAM_d_2  TTIENPLWSS SGNKGQRWNE AHVNIYPITS FQLIFEGIRG PGIEGDIAID
NCAM_c_1  TTIENPLWSS SGNKGQRWNE AHVNIYPITS FQLIFEGIRG PGIEGDIAID
NCAM_c_2  TTIENPLWSS SGNKGQRWNE AHVNIYPITS FQLIFEGIRG PGIEGDIAID

      901                                                       946
NCAM_d_1  DVSIAEGECA KQDLATKNSV DGAVGILVHI WLFPIIVLIS ILSPRR
NCAM_d_2  DVSIAEGECA KQDLATKNSV DGAVGILVHI WLFPIIVLIS ILSPRR
NCAM_c_1  DVSIAEGECA KQDLATKNSV DGAVGILVHI WLFPIIVLIS ILSPRR
NCAM_c_2  DVSIAEGECA KQDLATKNSV DGAVGILVHI WLFPIIVLIS ILSPRR
```

Fig. 4 (Cont.[4])

1. Brain +
2. Brain -
3. Cerebellum +
4. Cerebellum -
5. spleen adenocarcinoma +
6. spleen adenocarcinoma -
7. Thymus +
8. Thymus -
9. Lymphoblast +
10. Lymphoblast -
11. hela cells A +
12. hela cells A -
13. Heart +
14. Heart -



+ superscript reverse transcriptase

- superscript reverse transcriptase (control)

Fig. 5

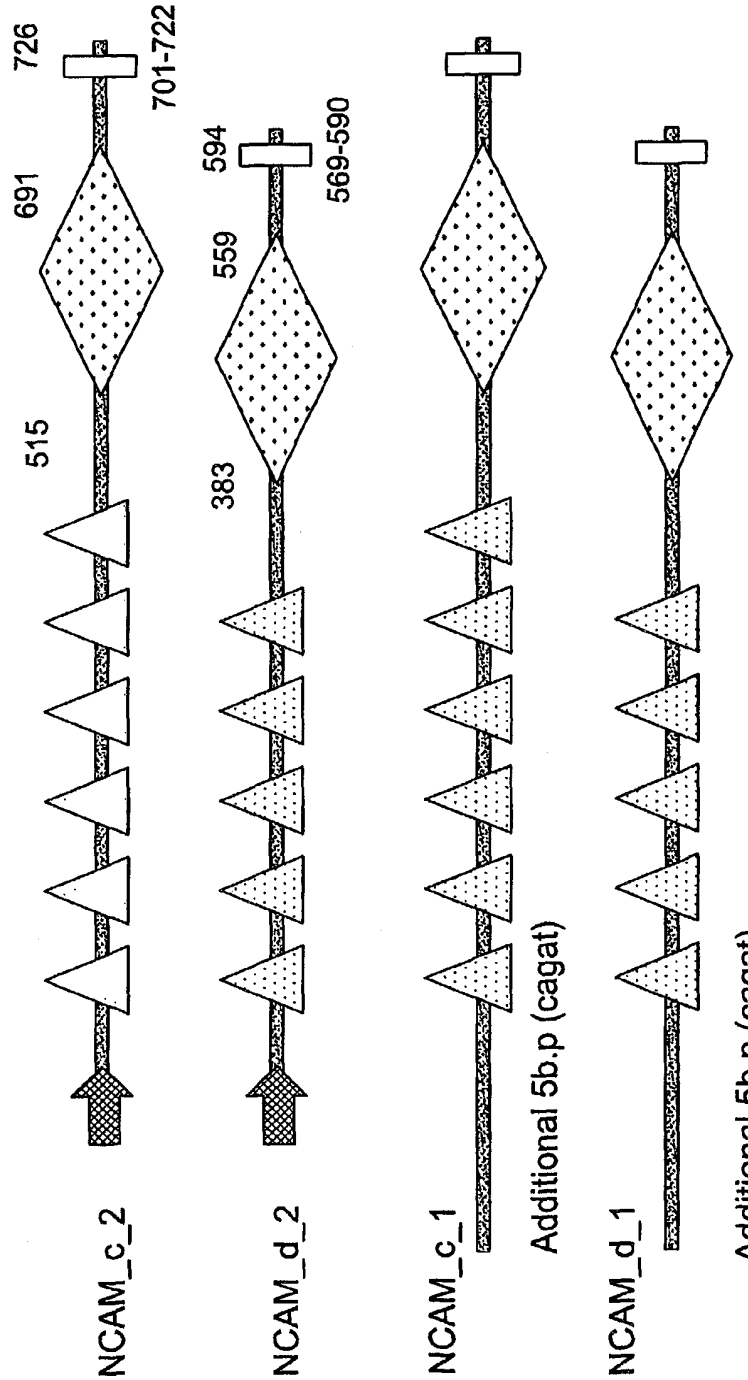

NCAM_c_2
NCAM_d_2
NCAM_c_1
Additional 5b.p (cagat)
NCAM_d_1
Additional 5b.p (cagat)
515
691
726
701-722
383
559
594
569-590
Met
Ig-domain
MAM domain
TM domain

Fig. 6

SEQUENCE LISTING

<110> COMPUGEN LTD

<120> N-CAM HOMOLOG

<130> 1289099-COMPUGEN

<140>
<141>

<150> 134932
<151> 2000-03-07

<160> 13

<170> PatentIn Ver. 2.1

<210> 1
<211> 3065
<212> DNA
<213> Homo sapiens

<400> 1
gtttttacc aacaaggcag aggatcagag atgtcttgct tcagctcctc ccacggttcg 60
tattgtgcac tcaggcttgg cctgtaacat tgaagaggag cgctactccg aaagggtcta 120
cactatccgg gaaggagaaa cccttgagtt gacttgtcta gtcacaggac atccacgtcc 180
acagatcagg tggaccaaaa cagcaggaag tgcctctgac agattccaag actcaagtgt 240
cttcaatgag actttgagga ttacaaatat tcagcgacac caaggaggcc ggtattactg 300
taaagcagag aatggcttgg ggtctccagc gataaagtca atcagagtgg atgtatacta 360
tttggatgat ccagtagtaa ctgttcatca aagtataggt gaagctaaag aacaatttta 420
ctatgagaga acagtgttcc tccggtgtgt tgccaattcc aatcctcctg ttcggtatag 480
ctggagacgt ggccaggagg tcttgctgca aggatctgat aaaggagttg agatttatga 540
accattcttt acccagggtg aaacaaagat cttaaaacta aagaatcttc gacctcagga 600
ctatgctaat tatagctgca ttgcttcagt gaggaatgta tgtaatattc ctgataagat 660
ggtgtcgttt agactgtcca ataaaacagc atcaccgtca attaaactct tggtggatga 720
tcctatagtt gtaaatcctg gagaggccat aacattagta tgtgttacaa caggaggaga 780
gcctgcacct tctctcacct gggtcaggtc ctttgggact ctgcctgaaa agactgtttt 840
gaatggagga actttgacca tacctgccat cacctcagat gatgctggta cttacagctg 900
cattgccaat aataatgtgg gaaaccctgc aaaaaagtcc accaacatca ttgtgagagc 960
attaaaaaaa ggacgatttt ggatcacacc agatccttat cacaaagatg acaacatcca 1020
gattggccgt gaggtgaaaa tatcttgcca agtagaagct gttccttctg aggagctaac 1080
atttagttgg tttaaaaatg gtcgtccatt aagaagttct gagcggatgg tcattacaca 1140
gactgatcct gatgtctctc cgggaacaac aaacttggac atcattgatt taaaattcac 1200
ggattctggg acgtacacat gtgtagcatc tctgaaggga ggaggaatat ctgatatcag 1260
tatcgatgtt aatatatcca gcagcacagt tccacccaat ctgactgttc cacaggaaaa 1320
atcaccattg gtcaccagag aaggagacac aatagaactg caatgtcaag taactggcaa 1380
acctaaacca atcatccttt ggtctagagc ggataaagaa gttgcaatgc ctgatggatc 1440
aatgcaaatg gagagttatg atggaacact gaggattgtg aatgtatcta gggaaatgtc 1500
aggaatgtac agatgtcaga ccagccaata caatggattt aacgtgaaac caagggaagc 1560
cttggtgcag ctcatcgttc agtatccccc tgcagtggaa ccagcattct tggaaatccg 1620
gcaaggacag gatcgaagtg tcactatgag ttgcagagta ctgagagcct atccaatacg 1680
ggtgctgacc tatgagtggc gcttgggcaa taaattatta cggacgggtc aatttgactc 1740
tcaggaatac acagagtacg ctgtgaagag tctttccaat gaaaactatg gggtttataa 1800
ctgtagcatc ataaatgaag ctggagctgg gagatgcagc tttcttgtta caggaaaggc 1860
ctatgctcca gaattctatt atgataccta caatccagta tggcagaaca gacaccgtgt 1920
ttattcttac agtctacagt ggacacagat gaatcctgat gcagtggatc ggattgttgc 1980
ataccggttg ggcatcaggc aggctggaca gcagcgctgg tgggagcagg agattaaaat 2040
aaatgggaat attcaaaagg gagaattaat tacatataac ttgacagagc taattaaacc 2100
agaagcttat gaagtccgac tgactcctct caccaaattt ggtgaaggag attcaacaat 2160
tcgtgtgatc aaatatagtg ctcctgtaaa tcctcatttg agagaatttc attgtggatt 2220

tgaagatggt aatatttgtt tgttcactca agatgataca gataattttg actggacaaa 2280
gcaaagtaca gcaacaagaa atacaaaata tactcctaat acaggaccta atgctgaccg 2340
tagtggctcc aaagaaggtt tttatatgta cattgagaca tcacgaccca gattggaagg 2400
cgaaaaggct cgacttctca gccctgtttt cagcatagct cccaaaaacc cttatggacc 2460
cacaaacact gcatattgtt tcagcttctt ttatcacatg tatggacaac atataggtgt 2520
cttaaatgtt tatctacgtt tgaaagggca aacaacaata gagaatccac tgtggtcttc 2580
aagtgggaat aaaggacaaa gatggaatga ggctcatgtt aatatatacc caattacttc 2640
atttcagctc atttttgaag gtatccgagg tcctggaata gaaggtgaca ttgctattga 2700
tgatgtatca attgcagaag gagaatgtgc aaaacaagac ctagcaacta agaattccgt 2760
tgatggtgct gttgggattt tggttcatat atggcttttt cccattatcg tcctcatctc 2820
tatcttaagt cctcgaaggt gaccttatcc tggcagaggc tataaaagat tcaccaggca 2880
ctggcatgaa gaaagagtct ttgtaaatgg acattgaaca aacaaactac caaagattcc 2940
tccactgact acnggactca aaaataaaat aataaaaaca aatttttta agcacctggg 3000
gataaaaaga catcatggaa gtataactta ttncagacta aacataaang ataatcgttg 3060
acctg 3065

<210> 2
<211> 2771
<212> DNA
<213> Homo sapiens

<400> 2
gttttttacc aacaaggcag aggatcagag atgtcttgct tcagctcctc ccacggttcg 60
tattgtgcac tcaggcttgg cctgtaacat tgaagaggag cgctactccg aaagggtcta 120
cactatccgg gaaggagaaa cccttgagtt gacttgtcta gtcacaggac atccacgtcc 180
acagatcagg tggaccaaaa cagcaggaag tgcctctgac agattccaag actcaagtgt 240
cttcaatgag actttgagga ttacaaatat tcagcgacac caaggaggcc ggtattactg 300
taaagcagag aatggcttgg ggtctccagc gataaagtca atcagagtgg atgtatacta 360
tttggatgat ccagtagtaa ctgttcatca aagtataggt gaagctaaag aacaatttta 420
ctatgagaga acagtgttcc tccggtgtgt tgccaattcc aatcctcctg ttaggtatag 480
ctggagacgt ggccaggagg tcttgctgca aggatctgat aaaggagttg agatttatga 540
accattcttt acccagggtg aaacaaagat cttaaaacta aagaatcttc gacctcagga 600
ctatgctaat tatagctgca ttgcttcagt gaggaatgta tgtaatattc ctgataagat 660
ggtgtcgttt agactgtcca ataaaacagc atcaccgtca attaaactct tggtggatga 720
tcctatagtt gtaaatcctg gagaggccat aacattagta tgtgttacaa caggaggaga 780
gcctgcacct tctctcacct gggtcaggtc ctttgggact ctgcctgaaa agactgtttt 840
gaatggagga actttgacca tacctgccat cacctcagat gatgctggta cttacagctg 900
cattgccaat aataatgtgg gaaaccctgc aaaaaagtcc accaacatca ttgtgagagc 960
attaaaaaaa ggtcgatttt ggatcacacc agatccttat cacaaagatg acaacatcca 1020
gattggccgt gaggtgaaaa tatcttgcca agtagaagct gttccttctg aggagctaac 1080
atttagttgg tttaaaaatg gtcgtccatt aagaagttct gagcggatgg tcattacaca 1140
gactgatcct gatgtctctc cgggaacaac aaacttggac atcattgatt taaaattcac 1200
ggattctggg acgtacacat gtgtagcatc tctgaaggga ggaggaatat ctgatatcag 1260
tatcgatgtt aatatatcca gcagcacagt tccacctgca gtggaaccag cattcttgga 1320
aatccggcaa ggacaggatc gaagtgtcac tatgagttgc agagtactga gagcctatcc 1380
aatacgggtg ctgacctatg agtggcgctt gggcaataaa ttattacgga cgggtcaatt 1440
tgactctcag gaatacacag agtacgctgt gaagagtctt tccaatgaaa actatggggt 1500
ttataactgt agcatcataa atgaagctgg agctgggaga tgcagctttc ttgttacagg 1560
aaaggcctat gctccagaat tctattatga tacctacaat ccagtatggc agaacagaca 1620
ccgtgtttat tcttacagtc tacagtggac acagatgaat cctgatgcag tggatcggat 1680
tgttgcatac cggttgggca tcaggcaggc tggacagcag cgctggtggg agcaggagat 1740
taaaataaat gggaatattc aaaagggaga attaattaca tataacttga cagagctaat 1800
taaaccagaa gcttatgaag tccgactgac tcctctcacc aaatttggtg aaggagattc 1860
aacaattcgt gtgatcaaat atagtgctcc tgtaaatcct catttgagag aatttcattg 1920
tggatttgaa gatggtaata tttgtttgtt cactcaagat gatacagata attttgactg 1980
gacaaagcaa agtacagcaa caagaaatac aaaatatact cctaatacag gacctaatgc 2040
tgaccgtagt ggctccaaag aaggttttta tatgtacatt gagacatcac gacccagatt 2100
ggaaggcgaa aaggctcgac ttctcagccc tgttttcagc atagctccca aaaaccctta 2160
tggacccaca aacactgcat attgtttcag cttcttttat cacatgtatg gacaacatat 2220
aggtgtctta aatgtttatc tacgtttgaa agggcaaaca acaatagaga atccactgtg 2280

gtcttcaagt gggaataaag gacaaagatg gaatgaggct catgttaata tatacccaat 2340
tacttcattt cagctcattt ttgaaggtat ccgaggtcct ggaatagaag gtgacattgc 2400
tattgatgat gtatcaattg cagaaggaga atgtgcaaaa caagacctag caactaagaa 2460
ttccgttgat ggtgctgttg ggattttggt tcatatatgg ctttttccca ttatcgtcct 2520
catctctatc ttaagtcctc gaaggtgacc ttatcctggc agaggctata aaagattcac 2580
caggcactgg catgaagaaa gagtctttgt aaatggacat tgaacaaaca aactaccaaa 2640
gattcctcca ctgactacng gactcaaaaa taaaataata aaaacaaatt tttttaagca 2700
cctggggata aaaagacatc atggaagtat aacttattnc agactaaaca taaangataa 2760
tcgttgacct g 2771

<210> 3
<211> 2766
<212> DNA
<213> Homo sapiens

<400> 3
gtttttttacc aacaaggcag aggatcagag atgtcttgct tcagctcctc ccacggttcg 60
tattgtgcac tcaggcttgg cctgtaacat tgaagaggag cgctactccg aaagggtcta 120
cactatccgg gaaggagaaa cccttgagtt gacttgtcta gtcacaggac atccacgtcc 180
acaggtggac caaaacagca ggaagtgcct ctgacagatt ccaagactca agtgtcttca 240
atgagacttt gaggattaca aatattcagc gacaccaagg aggcaggtat tactgtaaag 300
cagagaatgg cttggggtct ccagcgataa agtcaatcag agtggatgta tactatttgg 360
atgatccagt agtaactgtt catcaaagta taggtgaagc taaagaacaa ttttactatg 420
agagaacagt gttcctccgg tgtgttgcca attccaatcc tcctgttcgg tatagctgga 480
gacgtggcca ggaggtcttg ctgcaaggat ctgataaagg agttgagatt tatgaaccat 540
tctttaccca gggtgaaaca aagatcttaa aactaaagaa tcttcgacct caggactatg 600
ctaattatag ctgcattgct tcagtgagga atgtatgtaa tattcctgat aagatggtgt 660
cgtttagact gtccaataaa acagcatcac cgtcaattaa actcttggtg gatgatccta 720
tagttgtaaa tcctggagag gccataacat tagtatgtgt tacaacagga ggagagcctg 780
caccttctct cacctgggtc aggtcctttg ggactctgcc tgaaaagact gttttgaatg 840
gaggaacttt gaccatacct gccatcacct cagatgatgc tggtacttac agctgcattg 900
ccaataataa tgtgggaaac cctgcaaaaa agtccaccaa catcattgtg agagcattaa 960
aaaaaggacg attttggatc acaccagatc cttatcacaa agatgacaac atccagattg 1020
gccgtgaggt gaaaatatct tgccaagtag aagctgttcc ttctgaggag ctaacattta 1080
gttggtttaa aaatggtcgt ccattaagaa gttctgagcg gatggtcatt acacagactg 1140
atcctgatgt ctctccggga acaacaaact tggacatcat tgatttaaaa ttcacggatt 1200
ctgggacgta cacatgtgta gcatctctga agggaggagg aatatctgat atcagtatcg 1260
atgttaatat atccagcagc acagttccac ctgcagtgga accagcattc ttggaaatcc 1320
ggcaaggaca ggatcgaagt gtcactatga gttgcagagt actgagagcc tatccaatac 1380
gggtgctgac ctatgagtgg cgcttgggca ataaattatt acggacgggt caatttgact 1440
ctcaggaata cacagagtac gctgtgaaga gtctttccaa tgaaaactat ggggtttata 1500
actgtagcat cataaatgaa gctggagctg ggagatgcag ctttcttgtt acaggaaagg 1560
cctatgctcc agaattctat tatgatacct acaatccagt atggcagaac agacaccgtg 1620
tttattctta cagtctacag tggacacaga tgaatcctga tgcagtggat cggattgttg 1680
cataccggtt gggcatcagg caggctggac agcagcgctg gtgggagcag gagattaaaa 1740
tatatgggaa tattcaaaag ggagaattaa ttacatataa cttgacagag ctaattaaac 1800
cagaagctta tgaagtccga ctgactcctc tcaccaaatt tggtgaagga gattcaacaa 1860
ttcgtgtgat caaatatagt gctcctgtaa atcctcattt gagagaattt cattgtggat 1920
ttgaagatgg taatatttgt ttgttcactc aagatgatac agataatttt gactggacaa 1980
agcaaagtac agcaacaaga aatacaaaat atactcctaa tacaggacct aatgctgacc 2040
gtagtggctc caaagaaggt ttttatatgt acattgagac atcacgaccc agattggaag 2100
gcgaaaaggc tcgacttctc agccctgttt tcagcatagc tcccaaaaac ccttatggac 2160
ccacaaacac tgcatattgt ttcagcttct tttatcacat gtatggacaa catataggtg 2220
tcttaaatgt ttatctacgt ttgaaagggc aaacaacaat agagaatcca ctgtggtctt 2280
caagtgggaa taaaggacaa agatggaatg aggctcatgt taatatatac ccaattactt 2340
catttcagct catttttgaa ggtatccgag gtcctggaat agaaggtgac attgctattg 2400
atgatgtatc aattgcagaa ggagaatgtg caaaacaaga cctagcaact aagaattccg 2460
ttgatggtgc tgttgggatt ttggttcata tatggctttt tcccattatc gtcctcatct 2520
ctatcttaag tcctcgaagg tgaccttatc ctggcagagg ctataaaaga ttcaccaggc 2580
actggcatga agaaagagtc tttgtaaatg gacattgaac aaacaaacta ccaaagattc 2640

ctccactgac tacnggactc aaaaataaaa taataaaaac aaatttttt aagcacctgg 2700
ggataaaaag acatcatgga agtataactt attncagact aaacataaan gataatcgtt 2760
gacctg 2766

<210> 4
<211> 2771
<212> DNA
<213> Homo sapiens

<400> 4
gttttttacc aacaaggcag aggatcagag atgtcttgct tcagctcctc ccacggttcg 60
tattgtgcac tcaggcttgg cctgtaacat tgaagaggag cgctactccg aaagggtcta 120
cactatccgg gaaggagaaa cccttgagtt gacttgtcta gtcacaggac atccacgtcc 180
acagatcagg tggaccaaaa cagcaggaag tgcctctgac agattccaag actcaagtgt 240
cttcaatgag actttgagga ttacaaatat tcagcgacac caaggaggcc ggtattactg 300
taaagcagag aatggcttgg ggtctccagc gataaagtca atcagagtgg atgtatacta 360
tttggatgat ccagtagtaa ctgttcatca aagtataggt gaagctaaag aacaatttta 420
ctatgagaga acagtgttcc tccggtgtgt tgccaattcc aatcctcctg ttcggtatag 480
ctggagacgt ggccaggagg tcttgctgca aggatctgat aaaggagttg agatttatga 540
accattcttt acccagggtg aaacaaagat cttaaaacta aagaatcttc gacctcagga 600
ctatgctaat tatagctgca ttgcttcagt gaggaatgta tgtaatattc ctgataagat 660
ggtgtcgttt agactgtcca ataaaacagc atcaccgtca attaaactct tggtggatga 720
tcctatagtt gtaaatcctg gagaggccat aacattagta tgtgttacaa caggaggaga 780
gcctgcacct tctctcacct gggtcaggtc ctttgggact ctgcctgaaa agactgtttt 840
gaatggagga actttgacca tacctgccat cacctcagat gatgctggta cttacagctg 900
cattgccaat aataatgtgg gaaaccctgc aaaaaagtcc accaacatca ttgtgagagc 960
attaaaaaaa ggacgatttt ggatcacacc agatccttat cacaaagatg acaacatcca 1020
gattggccgt gaggtgaaaa tatcttgcca agtagaagct gttccttctg aggagctaac 1080
atttagttgg tttaaaaatg gtcgtccatt aagaagttct gagcggatgg tcattacaca 1140
gactgatcct gatgtctctc cgggaacaac aaacttggac atcattgatt taaaattcac 1200
ggattctggg acgtacacat gtgtagcatc tctgaaggga ggaggaatat ctgatatcag 1260
tatcgatgtt aatatatcca gcagcacagt tccacctgca gtggaaccag cattcttgga 1320
aatccggcaa ggacaggatc gaagtgtcac tatgagttgc agagtactga gagcctatcc 1380
aatacgggtg ctgacctatg agtggcgctt gggcaataaa ttattacgga cgggtcaatt 1440
tgactctcag gaatacacag agtacgctgt gaagagtctt tccaatgaaa actatggggt 1500
ttataactgt agcatcataa atgaagctgg agctgggaga tgcagctttc ttgttacagg 1560
aaaggcctat gctccagaat tctattatga tacctacaat ccagtatggc agaacagaca 1620
ccgtgtttat tcttacagtc tacagtggac acagatgaat cctgatgcag tggatcggat 1680
tgttgcatac cggttgggca tcaggcaggc tggacagcag cgctggtggg agcaggagat 1740
taaaatatat gggaatattc aaaagggaga attaattaca tataacttga cagagctaat 1800
taaaccagaa gcttatgaag tccgactgac tcctctcacc aaatttggtg aaggagattc 1860
aacaattcgt gtgatcaaat atagtgctcc tgtaaatcct catttgagag aatttcattg 1920
tggatttgaa gatggtaata tttgtttgtt cactcaagat gatacagata attttgactg 1980
gacaaagcaa agtacagcaa caagaaatac aaaatatact cctaatacag gacctaatgc 2040
tgaccgtagt ggctccaaag aaggttttta tatgtacatt gagacatcac gacccagatt 2100
ggaaggcgaa aaggctcgac ttctcagccc tgttttcagc atagctccca aaaaccctta 2160
tggacccaca aacactgcat attgtttcag cttcttttat cacatgtatg gacaacatat 2220
aggtgtctta aatgtttatc tacgtttgaa agggcaaaca acaatagaga atccactgtg 2280
gtcttcaagt gggaataaag gacaaagatg gaatgaggct catgttaata tatacccaat 2340
tacttcattt cagctcattt ttgaaggtat ccgaggtcct ggaatagaag gtgacattgc 2400
tattgatgat gtatcaattg cagaaggaga atgtgcaaaa caagacctag caactaagaa 2460
ttccgttgat ggtgctgttg ggattttggt tcatatatgg ctttttccca ttatcgtcct 2520
catctctatc ttaagtcctc gaaggtgacc ttatcctggc agaggctata aaagattcac 2580
caggcactgg catgaagaaa gagtctttgt aaatggacat tgaacaaaca aactaccaaa 2640
gattcctcca ctgactacng gactcaaaaa taaaataata aaaacaaatt tttttaagca 2700
cctggggata aaaagacatc atggaagtat aacttattnc agactaaaca taaangataa 2760
tcgttgacct g 2771

<210> 5

<211> 3060
<212> DNA
<213> Homo sapiens

<400> 5
gttttttacc aacaaggcag aggatcagag atgtcttgct tcagctcctc ccacggttcg 60
tattgtgcac tcaggcttgg cctgtaacat tgaagaggag cgctactccg aaagggtcta 120
cactatccgg gaaggagaaa cccttgagtt gacttgtcta gtcacaggac atccacgtcc 180
acaggtggac caaaacagca ggaagtgcct ctgacagatt ccaagactca agtgtcttca 240
atgagacttt gaggattaca aatattcagc gacaccaagg aggccggtat tactgtaaag 300
cagagaatgg cttggggtct ccagcgataa agtcaatcag agtggatgta tactatttgg 360
atgatccagt agtaactgtt catcaaagta taggtgaagc taaagaacaa ttttactatg 420
agagaacagt gttcctccgg tgtgttgcca attccaatcc tcctgttcgg tatagctgga 480
gacgtggcca ggaggtcttg ctgcaaggat ctgataaagg agttgagatt tatgaaccat 540
tctttaccca gggtgaaaca aagatcttaa aactaaagaa tcttcgacct caggactatg 600
ctaattatag ctgcattgct tcagtgagga atgtatgtaa tattcctgat aagatggtgt 660
cgtttagact gtccaataaa acagcatcac cgtcaattaa actcttggtg gatgatccta 720
tagttgtaaa tcctggagag gccataacat tagtatgtgt tacaacagga ggagagcctg 780
caccttctct cacctgggtc aggtcctttg ggactctgcc tgaaaagact gttttgaatg 840
gaggaacttt gaccatacct gccatcacct cagatgatgc tggtacttac agctgcattg 900
ccaataataa tgtgggaaac cctgcaaaaa agtccaccaa catcattgtg agagcattaa 960
aaaaaggacg attttggatc acaccagatc cttatcacaa agatgacaac atccagattg 1020
gccgtgaggt gaaaatatct tgccaagtag aagctgttcc ttctgaggag ctaacattta 1080
gttggtttaa aaatggtcgt ccattaagaa gttctgagcg gatggtcatt acacagactg 1140
atcctgatgt ctctccggga acaacaaact tggacatcat tgatttaaaa ttcacggatt 1200
ctgggacgta cacatgtgta gcatctctga agggaggagg aatatctgat atcagtatcg 1260
atgttaatat atccagcagc acagttccac ccaatctgac tgttccacag gaaaaatcac 1320
cattggtcac cagagaagga gacacaatag aactgcaatg tcaagtaact ggcaaaccta 1380
aaccaatcat cctttggtct agagcggata aagaagttgc aatgcctgat ggatcaatgc 1440
aaatggagag ttatgatgga acactgagga ttgtgaatgt atctagggaa atgtcaggaa 1500
tgtacagatg tcagaccagc caatacaatg gatttaacgt gaaaccaagg gaagccttgg 1560
tgcagctcat cgttcagtat ccccctgcag tggaaccagc attcttggaa atccggcaag 1620
gacaggatcg aagtgtcact atgagttgca gagtactgag agcctatcca atacgggtgc 1680
tgacctatga gtggcgcttg ggcaataaat tattacggac gggtcaattt gactctcagg 1740
aatacacaga gtacgctgtg aagagtcttt ccaatgaaaa ctatggggtt tataactgta 1800
gcatcataaa tgaagctgga gctgggagat gcagctttct tgttacagga aaggcctatg 1860
ctccagaatt ctattatgat acctacaatc cagtatggca gaacagacac cgtgtttatt 1920
cttacagtct acagtggaca cagatgaatc ctgatgcagt ggatcggatt gttgcatacc 1980
ggttgggcat caggcaggct ggacagcagc gctggtggga gcaggagatt aaaataaatg 2040
ggaatattca aaagggagaa ttaattacat ataacttgac agagctaatt aaaccagaag 2100
cttatgaagt ccgactgact cctctcacca aatttggtga aggagattca acaattcgtg 2160
tgatcaaata tagtgctcct gtaaatcctc atttgagaga atttcattgt ggatttgaag 2220
atggtaatat ttgtttgttc actcaagatg atacagataa ttttgactgg acaaagcaaa 2280
gtacagcaac aagaaataca aaatatactc ctaatacagg acctaatgct gaccgtagtg 2340
gctccaaaga aggtttttat atgtacattg agacatcacg acccagattg gaaggcgaaa 2400
aggctcgact tctcagccct gttttcagca tagctcccaa aaacccttat ggacccacaa 2460
acactgcata ttgtttcagc ttcttttatc acatgtatgg acaacatata ggtgtcttaa 2520
atgtttatct acgtttgaaa gggcaaacaa caatagagaa tccactgtgg tcttcaagtg 2580
ggaataaagg acaaagatgg aatgaggctc atgttaatat atacccaatt acttcatttc 2640
agctcatttt tgaaggtatc cgaggtcctg gaatagaagg tgacattgct attgatgatg 2700
tatcaattgc agaaggagaa tgtgcaaaac aagacctagc aactaagaat tccgttgatg 2760
gtgctgttgg gattttggtt catatatggc tttttcccat tatcgtcctc atctctatct 2820
taagtcctcg aaggtgacct tatcctggca gaggctataa aagattcacc aggcactggc 2880
atgaagaaag agtctttgta aatggacatt gaacaaacaa actaccaaag attcctccac 2940
tgactacngg actcaaaaat aaaataataa aaacaaattt ttttaagcac ctggggataa 3000
aaagacatca tggaagtata acttattnca gactaaacat aaangataat cgttgacctg 3060

<210> 6
<211> 3065
<212> DNA

<213> Homo sapiens

<400> 6
gttttttacc aacaaggcag aggatcagag atgtcttgct tcagctcctc ccacggttcg 60
tattgtgcac tcaggcttgg cctgtaacat tgaagaggag cgctactccg aaagggtcta 120
cactatccgg gaaggagaaa cccttgagtt gacttgtcta gtcacaggac atccacgtcc 180
acagatcagg tggaccaaaa cagcaggaag tgcctctgac agattccaag actcaagtgt 240
cttcaatgag actttgagga ttacaaatat tcagcgacac caaggaggcc ggtattactg 300
taaagcagag aatggcttgg ggtctccagc gataaagtca atcagagtgg atgtatacta 360
tttggatgat ccagtagtaa ctgttcatca aagtataggt gaagctaaag aacaatttta 420
ctatgagaga acagtgttcc tccggtgtgt tgccaattcc aatcctcctg ttcggtatag 480
ctggagacgt ggccaggagg tcttgctgca aggatctgat aaaggagttg agatttatga 540
accattcttt acccagggtg aaacaaagat cttaaaacta aagaatcttc gacctcagga 600
ctatgctaat tatagctgca ttgcttcagt gaggaatgta tgtaatattc ctgataagat 660
ggtgtcgttt agactgtcca ataaaacagc atcaccgtca attaaactct tggtggatga 720
tcctatagtt gtaaatcctg gagaggccat aacattagta tgtgttacaa caggaggaga 780
gcctgcacct tctctcacct gggtcaggtc ctttgggact ctgcctgaaa agactgtttt 840
gaatggagga actttgacca tacctgccat cacctcagat gatgctggta cttacagctg 900
cattgccaat aataatgtgg gaaaccctgc aaaaaagtcc accaacatca ttgtgagagc 960
attaaaaaaa ggacgatttt ggatcacacc agatccttat cacaaagatg acaacatcca 1020
gattggccgt gaggtgaaaa tatcttgcca agtagaagct gttccttctg aggagctaac 1080
atttagttgg tttaaaaatg gtcgtccatt aagaagttct gagcggatgg tcattacaca 1140
gactgatcct gatgtctctc cgggaacaac aaacttggac atcattgatt taaaattcac 1200
ggattctggg acgtacacat gtgtagcatc tctgaaggga ggaggaatat ctgatatcag 1260
tatcgatgtt aatatatcca gcagcacagt tccacccaat ctgactgttc cacaggaaaa 1320
atcaccattg gtcaccagag aaggagacac aatagaactg caatgtcaag taactggcaa 1380
acctaaacca atcatccttt ggtctagagc ggataaagaa gttgcaatgc ctgatggatc 1440
aatgcaaatg gagagttatg atggaacact gaggattgtg aatgtatcta gggaaatgtc 1500
aggaatgtac agatgtcaga ccagccaata caatggattt aacgtgaaac caagggaagc 1560
cttggtgcag ctcatcgttc agtatccccc tgcagtggaa ccagcattct tggaaatccg 1620
gcaaggacag gatcgaagtg tcactatgag ttgcagagta ctgagagcct atccaatacg 1680
ggtgctgacc tatgagtggc gcttgggcaa taaattatta cggacggtc aatttgactc 1740
tcaggaatac acagagtacg ctgtgaagag tctttccaat gaaaactatg gggtttataa 1800
ctgtagcatc ataaatgaag ctggagctgg gagatgcagc tttcttgtta caggaaaggc 1860
ctatgctcca gaattctatt atgataccta caatccagta tggcagaaca gacaccgtgt 1920
ttattcttac agtctacagt ggacacagat gaatcctgat gcagtggatc ggattgttgc 1980
ataccggttg ggcatcaggc aggctggaca gcagcgctgg tgggagcagg agattaaaat 2040
aaatgggaat attcaaaagg gagaattaat tacatataac ttgacagagc taattaaacc 2100
agaagcttat gaagtccgac tgactcctct caccaaattt ggtgaaggag attcaacaat 2160
tcgtgtgatc aaatatagtg ctcctgtaaa tcctcatttg agagaatttc attgtggatt 2220
tgaagatggt aatatttgtt tgttcactca agatgataca gataattttg actggacaaa 2280
gcaaagtaca gcaacaagaa atacaaaata tactcctaat acaggaccta atgctgaccg 2340
tagtggctcc aaagaaggtt tttatatgta cattgagaca tcacgaccca gattggaagg 2400
cgaaaaggct cgacttctca gccctgtttt cagcatagct cccaaaaacc cttatggacc 2460
cacaaacact gcatattgtt tcagcttctt ttatcacatg tatggacaac atataggtgt 2520
cttaaatgtt tatctacgtt tgaaagggca aacaacaata gagaatccac tgtggtcttc 2580
aagtgggaat aaaggacaaa gatggaatga ggctcatgtt aatatatacc caattacttc 2640
atttcagctc atttttgaag gtatccgagg tcctggaata gaaggtgaca ttgctattga 2700
tgatgtatca attgcagaag gagaatgtgc aaaacaagac ctagcaacta agaattccgt 2760
tgatggtgct gttgggattt tggttcatat atggcttttt cccattatcg tcctcatctc 2820
tatcttaagt cctcgaaggt gaccttatcc tggcagaggc tataaaagat tcaccaggca 2880
ctggcatgaa gaaagagtct ttgtaaatgg acattgaaca aacaaactac caaagattcc 2940
tccactgact acnggactca aaaataaaat aataaaaaca aatttttta agcacctggg 3000
gataaaaaga catcatggaa gtataactta ttncagacta aacataaang ataatcgttg 3060
acctg 3065

<210> 7
<211> 762
<212> DNA
<213> Homo sapiens

<400> 7
ccctgtttca gcatagctcc caaaaatcca tatggaccta caaatagtgc atattgtttc 60
agtttcttct atcacatgta cgggcaacat ataggggttt taaatgtata tctacgtttg 120
aaagggcaga caacgataga gaatccgcta tggtcgtcga gtgggaacaa aggacaacga 180
tggaatgaag ctcatgttaa tatatatcca attcttcatt tcagttaatt tttgaaggca 240
ttcgaggtcc tgggatagag ggtgacatcg ccattgatga tgtatcaatt gctgaaggag 300
aatgtgcaaa acaagaccta ccaactaaga attccgtgga tggtgctgtt gggatcttag 360
ttcatatatg gctttttcca gttatcatcc tcatctctat cttaagccct cgaaggtgac 420
cttatcctgg cagaggctat taatgattca ccaggcactg gcatgaagaa agaagctttg 480
taaatggaca ctgaaaaaac aaactaccaa agattcatcc actgactact gactcaaaaa 540
ttaaataata aaaacaacat ttttttaagc actggggatt aaaaaaaaag acatcatgga 600
aatataactt attccaaact acatatagaa gatattcttg gcctgggtac aaaagagccc 660
ttcaggtgct tttgtggcta aaatgattac agcttcatct ggttgaactc tggaaaaaga 720
aaataaaaga aaaagaacga ctgagagaaa aaaaaaaaaa aa 762

<210> 8
<211> 1018
<212> PRT
<213> Homo sapiens

<400> 8
Phe Phe Thr Asn Lys Ala Glu Asp Gln Arg Cys Leu Ala Ser Ala Pro
1 5 10 15

Pro Thr Val Arg Ile Val His Ser Gly Leu Ala Cys Asn Ile Glu Glu
20 25 30

Glu Arg Tyr Ser Glu Arg Val Tyr Thr Ile Arg Glu Gly Glu Thr Leu
35 40 45

Glu Leu Thr Cys Leu Val Thr Gly His Pro Arg Pro Gln Ile Arg Trp
50 55 60

Thr Lys Thr Ala Gly Ser Ala Ser Asp Arg Phe Gln Asp Ser Ser Val
65 70 75 80

Phe Asn Glu Thr Leu Arg Ile Thr Asn Ile Gln Arg His Gln Gly Gly
85 90 95

Arg Tyr Tyr Cys Lys Ala Glu Asn Gly Leu Gly Ser Pro Ala Ile Lys
100 105 110

Ser Ile Arg Val Asp Val Tyr Tyr Leu Asp Asp Pro Val Val Thr Val
115 120 125

His Gln Ser Ile Gly Glu Ala Lys Glu Gln Phe Tyr Tyr Glu Arg Thr
130 135 140

Val Phe Leu Arg Cys Val Ala Asn Ser Asn Pro Pro Val Arg Tyr Ser
145 150 155 160

Trp Arg Arg Gly Gln Glu Val Leu Leu Gln Gly Ser Asp Lys Gly Val
165 170 175

Glu Ile Tyr Glu Pro Phe Phe Thr Gln Gly Glu Thr Lys Ile Leu Lys
180 185 190

Leu Lys Asn Leu Arg Pro Gln Asp Tyr Ala Asn Tyr Ser Cys Ile Ala
195 200 205

```
Ser Val Arg Asn Val Cys Asn Ile Pro Asp Lys Met Val Ser Phe Arg
    210                 215                 220

Leu Ser Asn Lys Thr Ala Ser Pro Ser Ile Lys Leu Leu Val Asp Asp
225                 230                 235                 240

Pro Ile Val Val Asn Pro Gly Glu Ala Ile Thr Leu Val Cys Val Thr
                245                 250                 255

Thr Gly Gly Glu Pro Ala Pro Ser Leu Thr Trp Val Arg Ser Phe Gly
            260                 265                 270

Thr Leu Pro Glu Lys Thr Val Leu Asn Gly Gly Thr Leu Thr Ile Pro
        275                 280                 285

Ala Ile Thr Ser Asp Asp Ala Gly Thr Tyr Ser Cys Ile Ala Asn Asn
    290                 295                 300

Asn Val Gly Asn Pro Ala Lys Lys Ser Thr Asn Ile Ile Val Arg Ala
305                 310                 315                 320

Leu Lys Lys Gly Arg Phe Trp Ile Thr Pro Asp Pro Tyr His Lys Asp
                325                 330                 335

Asp Asn Ile Gln Ile Gly Arg Glu Val Lys Ile Ser Cys Gln Val Glu
            340                 345                 350

Ala Val Pro Ser Glu Glu Leu Thr Phe Ser Trp Phe Lys Asn Gly Arg
        355                 360                 365

Pro Leu Arg Ser Ser Glu Arg Met Val Ile Thr Gln Thr Asp Pro Asp
    370                 375                 380

Val Ser Pro Gly Thr Thr Asn Leu Asp Ile Ile Asp Leu Lys Phe Thr
385                 390                 395                 400

Asp Ser Gly Thr Tyr Thr Cys Val Ala Ser Leu Lys Gly Gly Gly Ile
                405                 410                 415

Ser Asp Ile Ser Ile Asp Val Asn Ile Ser Ser Ser Thr Val Pro Pro
            420                 425                 430

Asn Leu Thr Val Pro Gln Glu Lys Ser Pro Leu Val Thr Arg Glu Gly
        435                 440                 445

Asp Thr Ile Glu Leu Gln Cys Gln Val Thr Gly Lys Pro Lys Pro Ile
    450                 455                 460

Ile Leu Trp Ser Arg Ala Asp Lys Glu Val Ala Met Pro Asp Gly Ser
465                 470                 475                 480

Met Gln Met Glu Ser Tyr Asp Gly Thr Leu Arg Ile Val Asn Val Ser
                485                 490                 495

Arg Glu Met Ser Gly Met Tyr Arg Cys Gln Thr Ser Gln Tyr Asn Gly
            500                 505                 510

Phe Asn Val Lys Pro Arg Glu Ala Leu Val Gln Leu Ile Val Gln Tyr
        515                 520                 525
```

Pro Pro Ala Val Glu Pro Ala Phe Leu Glu Ile Arg Gln Gly Gln Asp
530 535 540

Arg Ser Val Thr Met Ser Cys Arg Val Leu Arg Ala Tyr Pro Ile Arg
545 550 555 560

Val Leu Thr Tyr Glu Trp Arg Leu Gly Asn Lys Leu Leu Arg Thr Gly
565 570 575

Gln Phe Asp Ser Gln Glu Tyr Thr Glu Tyr Ala Val Lys Ser Leu Ser
580 585 590

Asn Glu Asn Tyr Gly Val Tyr Asn Cys Ser Ile Ile Asn Glu Ala Gly
595 600 605

Ala Gly Arg Cys Ser Phe Leu Val Thr Gly Lys Ala Tyr Ala Pro Glu
610 615 620

Phe Tyr Tyr Asp Thr Tyr Asn Pro Val Trp Gln Asn Arg His Arg Val
625 630 635 640

Tyr Ser Tyr Ser Leu Gln Trp Thr Gln Met Asn Pro Asp Ala Val Asp
645 650 655

Arg Ile Val Ala Tyr Arg Leu Gly Ile Arg Gln Ala Gly Gln Gln Arg
660 665 670

Trp Trp Glu Gln Glu Ile Lys Ile Asn Gly Asn Ile Gln Lys Gly Glu
675 680 685

Leu Ile Thr Tyr Asn Leu Thr Glu Leu Ile Lys Pro Glu Ala Tyr Glu
690 695 700

Val Arg Leu Thr Pro Leu Thr Lys Phe Gly Glu Gly Asp Ser Thr Ile
705 710 715 720

Arg Val Ile Lys Tyr Ser Ala Pro Val Asn Pro His Leu Arg Glu Phe
725 730 735

His Cys Gly Phe Glu Asp Gly Asn Ile Cys Leu Phe Thr Gln Asp Asp
740 745 750

Thr Asp Asn Phe Asp Trp Thr Lys Gln Ser Thr Ala Thr Arg Asn Thr
755 760 765

Lys Tyr Thr Pro Asn Thr Gly Pro Asn Ala Asp Arg Ser Gly Ser Lys
770 775 780

Glu Gly Phe Tyr Met Tyr Ile Glu Thr Ser Arg Pro Arg Leu Glu Gly
785 790 795 800

Glu Lys Ala Arg Leu Leu Ser Pro Val Phe Ser Ile Ala Pro Lys Asn
805 810 815

Pro Tyr Gly Pro Thr Asn Thr Ala Tyr Cys Phe Ser Phe Phe Tyr His
820 825 830

Met Tyr Gly Gln His Ile Gly Val Leu Asn Val Tyr Leu Arg Leu Lys
835 840 845

Gly Gln Thr Thr Ile Glu Asn Pro Leu Trp Ser Ser Ser Gly Asn Lys

```
    850                 855                 860

Gly Gln Arg Trp Asn Glu Ala His Val Asn Ile Tyr Pro Ile Thr Ser
865                 870                 875                 880

Phe Gln Leu Ile Phe Glu Gly Ile Arg Gly Pro Gly Ile Glu Gly Asp
                885                 890                 895

Ile Ala Ile Asp Asp Val Ser Ile Ala Glu Gly Glu Cys Ala Lys Gln
            900                 905                 910

Asp Leu Ala Thr Lys Asn Ser Val Asp Gly Ala Val Gly Ile Leu Val
        915                 920                 925

His Ile Trp Leu Phe Pro Ile Ile Val Leu Ile Ser Ile Leu Ser Pro
    930                 935                 940

Arg Arg Pro Tyr Pro Gly Arg Gly Tyr Lys Arg Phe Thr Arg His Trp
945                 950                 955                 960

His Glu Glu Arg Val Phe Val Asn Gly His Thr Asn Lys Leu Pro Lys
                965                 970                 975

Ile Pro Pro Leu Thr Thr Gly Leu Lys Asn Lys Ile Ile Lys Thr Asn
            980                 985                 990

Phe Phe Lys His Leu Gly Ile Lys Arg His His Gly Ser Ile Thr Tyr
        995                1000                1005

Xaa Arg Leu Asn Ile Xaa Asp Asn Arg Pro
   1010                1015


<210> 9
<211> 920
<212> PRT
<213> Homo sapiens

<400> 9
Phe Phe Thr Asn Lys Ala Glu Asp Gln Arg Cys Leu Ala Ser Ala Pro
  1               5                  10                  15

Pro Thr Val Arg Ile Val His Ser Gly Leu Ala Cys Asn Ile Glu Glu
            20                  25                  30

Glu Arg Tyr Ser Glu Arg Val Tyr Thr Ile Arg Glu Gly Glu Thr Leu
        35                  40                  45

Glu Leu Thr Cys Leu Val Thr Gly His Pro Arg Pro Gln Ile Arg Trp
    50                  55                  60

Thr Lys Thr Ala Gly Ser Ala Ser Asp Arg Phe Gln Asp Ser Ser Val
65                  70                  75                  80

Phe Asn Glu Thr Leu Arg Ile Thr Asn Ile Gln Arg His Gln Gly Gly
                85                  90                  95

Arg Tyr Tyr Cys Lys Ala Glu Asn Gly Leu Gly Ser Pro Ala Ile Lys
            100                 105                 110

Ser Ile Arg Val Asp Val Tyr Tyr Leu Asp Asp Pro Val Val Thr Val
```

```
        115                 120                 125

His Gln Ser Ile Gly Glu Ala Lys Glu Gln Phe Tyr Tyr Glu Arg Thr
    130                 135                 140

Val Phe Leu Arg Cys Val Ala Asn Ser Asn Pro Pro Val Arg Tyr Ser
145                 150                 155                 160

Trp Arg Arg Gly Gln Glu Val Leu Leu Gln Gly Ser Asp Lys Gly Val
                165                 170                 175

Glu Ile Tyr Glu Pro Phe Phe Thr Gln Gly Glu Thr Lys Ile Leu Lys
            180                 185                 190

Leu Lys Asn Leu Arg Pro Gln Asp Tyr Ala Asn Tyr Ser Cys Ile Ala
        195                 200                 205

Ser Val Arg Asn Val Cys Asn Ile Pro Asp Lys Met Val Ser Phe Arg
    210                 215                 220

Leu Ser Asn Lys Thr Ala Ser Pro Ser Ile Lys Leu Leu Val Asp Asp
225                 230                 235                 240

Pro Ile Val Val Asn Pro Gly Glu Ala Ile Thr Leu Val Cys Val Thr
                245                 250                 255

Thr Gly Gly Glu Pro Ala Pro Ser Leu Thr Trp Val Arg Ser Phe Gly
            260                 265                 270

Thr Leu Pro Glu Lys Thr Val Leu Asn Gly Gly Thr Leu Thr Ile Pro
        275                 280                 285

Ala Ile Thr Ser Asp Asp Ala Gly Thr Tyr Ser Cys Ile Ala Asn Asn
    290                 295                 300

Asn Val Gly Asn Pro Ala Lys Lys Ser Thr Asn Ile Ile Val Arg Ala
305                 310                 315                 320

Leu Lys Lys Gly Arg Phe Trp Ile Thr Pro Asp Pro Tyr His Lys Asp
                325                 330                 335

Asp Asn Ile Gln Ile Gly Arg Glu Val Lys Ile Ser Cys Gln Val Glu
            340                 345                 350

Ala Val Pro Ser Glu Glu Leu Thr Phe Ser Trp Phe Lys Asn Gly Arg
        355                 360                 365

Pro Leu Arg Ser Ser Glu Arg Met Val Ile Thr Gln Thr Asp Pro Asp
    370                 375                 380

Val Ser Pro Gly Thr Thr Asn Leu Asp Ile Ile Asp Leu Lys Phe Thr
385                 390                 395                 400

Asp Ser Gly Thr Tyr Thr Cys Val Ala Ser Leu Lys Gly Gly Gly Ile
                405                 410                 415

Ser Asp Ile Ser Ile Asp Val Asn Ile Ser Ser Ser Thr Val Pro Pro
            420                 425                 430

Ala Val Glu Pro Ala Phe Leu Glu Ile Arg Gln Gly Gln Asp Arg Ser
        435                 440                 445
```

```
Val Thr Met Ser Cys Arg Val Leu Arg Ala Tyr Pro Ile Arg Val Leu
    450                 455                 460

Thr Tyr Glu Trp Arg Leu Gly Asn Lys Leu Leu Arg Thr Gly Gln Phe
465                 470                 475                 480

Asp Ser Gln Glu Tyr Thr Glu Tyr Ala Val Lys Ser Leu Ser Asn Glu
                485                 490                 495

Asn Tyr Gly Val Tyr Asn Cys Ser Ile Ile Asn Glu Ala Gly Ala Gly
            500                 505                 510

Arg Cys Ser Phe Leu Val Thr Gly Lys Ala Tyr Ala Pro Glu Phe Tyr
        515                 520                 525

Tyr Asp Thr Tyr Asn Pro Val Trp Gln Asn Arg His Arg Val Tyr Ser
    530                 535                 540

Tyr Ser Leu Gln Trp Thr Gln Met Asn Pro Asp Ala Val Asp Arg Ile
545                 550                 555                 560

Val Ala Tyr Arg Leu Gly Ile Arg Gln Ala Gly Gln Gln Arg Trp Trp
                565                 570                 575

Glu Gln Glu Ile Lys Ile Asn Gly Asn Ile Gln Lys Gly Glu Leu Ile
            580                 585                 590

Thr Tyr Asn Leu Thr Glu Leu Ile Lys Pro Glu Ala Tyr Glu Val Arg
        595                 600                 605

Leu Thr Pro Leu Thr Lys Phe Gly Glu Gly Asp Ser Thr Ile Arg Val
    610                 615                 620

Ile Lys Tyr Ser Ala Pro Val Asn Pro His Leu Arg Glu Phe His Cys
625                 630                 635                 640

Gly Phe Glu Asp Gly Asn Ile Cys Leu Phe Thr Gln Asp Asp Thr Asp
                645                 650                 655

Asn Phe Asp Trp Thr Lys Gln Ser Thr Ala Thr Arg Asn Thr Lys Tyr
            660                 665                 670

Thr Pro Asn Thr Gly Pro Asn Ala Asp Arg Ser Gly Ser Lys Glu Gly
        675                 680                 685

Phe Tyr Met Tyr Ile Glu Thr Ser Arg Pro Arg Leu Glu Gly Glu Lys
    690                 695                 700

Ala Arg Leu Leu Ser Pro Val Phe Ser Ile Ala Pro Lys Asn Pro Tyr
705                 710                 715                 720

Gly Pro Thr Asn Thr Ala Tyr Cys Phe Ser Phe Phe Tyr His Met Tyr
                725                 730                 735

Gly Gln His Ile Gly Val Leu Asn Val Tyr Leu Arg Leu Lys Gly Gln
            740                 745                 750

Thr Thr Ile Glu Asn Pro Leu Trp Ser Ser Ser Gly Asn Lys Gly Gln
        755                 760                 765
```

```
Arg Trp Asn Glu Ala His Val Asn Ile Tyr Pro Ile Thr Ser Phe Gln
    770                 775                 780

Leu Ile Phe Glu Gly Ile Arg Gly Pro Gly Ile Glu Gly Asp Ile Ala
785                 790                 795                 800

Ile Asp Asp Val Ser Ile Ala Glu Gly Glu Cys Ala Lys Gln Asp Leu
                805                 810                 815

Ala Thr Lys Asn Ser Val Asp Gly Ala Val Gly Ile Leu Val His Ile
            820                 825                 830

Trp Leu Phe Pro Ile Ile Val Leu Ile Ser Ile Leu Ser Pro Arg Arg
        835                 840                 845

Pro Tyr Pro Gly Arg Gly Tyr Lys Arg Phe Thr Arg His Trp His Glu
    850                 855                 860

Glu Arg Val Phe Val Asn Gly His Thr Asn Lys Leu Pro Lys Ile Pro
865                 870                 875                 880

Pro Leu Thr Thr Gly Leu Lys Asn Lys Ile Ile Lys Thr Asn Phe Phe
                885                 890                 895

Lys His Leu Gly Ile Lys Arg His His Gly Ser Ile Thr Tyr Xaa Arg
            900                 905                 910

Leu Asn Ile Xaa Asp Asn Arg Pro
        915                 920
```

<210> 10
<211> 848
<212> PRT
<213> Homo sapiens

<400> 10

```
Phe Phe Thr Asn Lys Ala Glu Asp Gln Arg Cys Leu Ala Ser Ala Pro
  1               5                  10                  15

Pro Thr Val Arg Ile Val His Ser Gly Leu Ala Cys Asn Ile Glu Glu
            20                  25                  30

Glu Arg Tyr Ser Glu Arg Val Tyr Thr Ile Arg Glu Gly Glu Thr Leu
        35                  40                  45

Glu Leu Thr Cys Leu Val Thr Gly His Pro Arg Pro Gln Ile Arg Trp
    50                  55                  60

Thr Lys Thr Ala Gly Ser Ala Ser Asp Arg Phe Gln Asp Ser Ser Val
65                  70                  75                  80

Phe Asn Glu Thr Leu Arg Ile Thr Asn Ile Gln Arg His Gln Gly Gly
                85                  90                  95

Arg Tyr Tyr Cys Lys Ala Glu Asn Gly Leu Gly Ser Pro Ala Ile Lys
            100                 105                 110

Ser Ile Arg Val Asp Val Tyr Tyr Leu Asp Asp Pro Val Val Thr Val
        115                 120                 125
```

```
His Gln Ser Ile Gly Glu Ala Lys Glu Gln Phe Tyr Tyr Glu Arg Thr
130                 135                 140

Val Phe Leu Arg Cys Val Ala Asn Ser Asn Pro Pro Val Arg Tyr Ser
145                 150                 155                 160

Trp Arg Arg Gly Gln Glu Val Leu Leu Gln Gly Ser Asp Lys Gly Val
                165                 170                 175

Glu Ile Tyr Glu Pro Phe Phe Thr Gln Gly Glu Thr Lys Ile Leu Lys
            180                 185                 190

Leu Lys Asn Leu Arg Pro Gln Asp Tyr Ala Asn Tyr Ser Cys Ile Ala
        195                 200                 205

Ser Val Arg Asn Val Cys Asn Ile Pro Asp Lys Met Val Ser Phe Arg
    210                 215                 220

Leu Ser Asn Lys Thr Ala Ser Pro Ser Ile Lys Leu Leu Val Asp Asp
225                 230                 235                 240

Pro Ile Val Val Asn Pro Gly Glu Ala Ile Thr Leu Val Cys Val Thr
                245                 250                 255

Thr Gly Gly Glu Pro Ala Pro Ser Leu Thr Trp Val Arg Ser Phe Gly
            260                 265                 270

Thr Leu Pro Glu Lys Thr Val Leu Asn Gly Gly Thr Leu Thr Ile Pro
        275                 280                 285

Ala Ile Thr Ser Asp Asp Ala Gly Thr Tyr Ser Cys Ile Ala Asn Asn
    290                 295                 300

Asn Val Gly Asn Pro Ala Lys Lys Ser Thr Asn Ile Ile Val Arg Ala
305                 310                 315                 320

Leu Lys Lys Gly Arg Phe Trp Ile Thr Pro Asp Pro Tyr His Lys Asp
                325                 330                 335

Asp Asn Ile Gln Ile Gly Arg Glu Val Lys Ile Ser Cys Gln Val Glu
            340                 345                 350

Ala Val Pro Ser Glu Glu Leu Thr Phe Ser Trp Phe Lys Asn Gly Arg
        355                 360                 365

Pro Leu Arg Ser Ser Glu Arg Met Val Ile Thr Gln Thr Asp Pro Asp
    370                 375                 380

Val Ser Pro Gly Thr Thr Asn Leu Asp Ile Ile Asp Leu Lys Phe Thr
385                 390                 395                 400

Asp Ser Gly Thr Tyr Thr Cys Val Ala Ser Leu Lys Gly Gly Gly Ile
                405                 410                 415

Ser Asp Ile Ser Ile Asp Val Asn Ile Ser Ser Ser Thr Val Pro Pro
            420                 425                 430

Ala Val Glu Pro Ala Phe Leu Glu Ile Arg Gln Gly Gln Asp Arg Ser
        435                 440                 445

Val Thr Met Ser Cys Arg Val Leu Arg Ala Tyr Pro Ile Arg Val Leu
```

```
    450                 455                 460

Thr Tyr Glu Trp Arg Leu Gly Asn Lys Leu Leu Arg Thr Gly Gln Phe
465                 470                 475                 480

Asp Ser Gln Glu Tyr Thr Glu Tyr Ala Val Lys Ser Leu Ser Asn Glu
                485                 490                 495

Asn Tyr Gly Val Tyr Asn Cys Ser Ile Ile Asn Glu Ala Gly Ala Gly
            500                 505                 510

Arg Cys Ser Phe Leu Val Thr Gly Lys Ala Tyr Ala Pro Glu Phe Tyr
        515                 520                 525

Tyr Asp Thr Tyr Asn Pro Val Trp Gln Asn Arg His Arg Val Tyr Ser
    530                 535                 540

Tyr Ser Leu Gln Trp Thr Gln Met Asn Pro Asp Ala Val Asp Arg Ile
545                 550                 555                 560

Val Ala Tyr Arg Leu Gly Ile Arg Gln Ala Gly Gln Gln Arg Trp Trp
                565                 570                 575

Glu Gln Glu Ile Lys Ile Tyr Gly Asn Ile Gln Lys Gly Glu Leu Ile
            580                 585                 590

Thr Tyr Asn Leu Thr Glu Leu Ile Lys Pro Glu Ala Tyr Glu Val Arg
        595                 600                 605

Leu Thr Pro Leu Thr Lys Phe Gly Glu Gly Asp Ser Thr Ile Arg Val
    610                 615                 620

Ile Lys Tyr Ser Ala Pro Val Asn Pro His Leu Arg Glu Phe His Cys
625                 630                 635                 640

Gly Phe Glu Asp Gly Asn Ile Cys Leu Phe Thr Gln Asp Asp Thr Asp
                645                 650                 655

Asn Phe Asp Trp Thr Lys Gln Ser Thr Ala Thr Arg Asn Thr Lys Tyr
            660                 665                 670

Thr Pro Asn Thr Gly Pro Asn Ala Asp Arg Ser Gly Ser Lys Glu Gly
        675                 680                 685

Phe Tyr Met Tyr Ile Glu Thr Ser Arg Pro Arg Leu Glu Gly Glu Lys
    690                 695                 700

Ala Arg Leu Leu Ser Pro Val Phe Ser Ile Ala Pro Lys Asn Pro Tyr
705                 710                 715                 720

Gly Pro Thr Asn Thr Ala Tyr Cys Phe Ser Phe Phe Tyr His Met Tyr
                725                 730                 735

Gly Gln His Ile Gly Val Leu Asn Val Tyr Leu Arg Leu Lys Gly Gln
            740                 745                 750

Thr Thr Ile Glu Asn Pro Leu Trp Ser Ser Ser Gly Asn Lys Gly Gln
        755                 760                 765

Arg Trp Asn Glu Ala His Val Asn Ile Tyr Pro Ile Thr Ser Phe Gln
    770                 775                 780
```

```
Leu Ile Phe Glu Gly Ile Arg Gly Pro Gly Ile Glu Gly Asp Ile Ala
785                 790                 795                 800

Ile Asp Asp Val Ser Ile Ala Glu Gly Glu Cys Ala Lys Gln Asp Leu
                805                 810                 815

Ala Thr Lys Asn Ser Val Asp Gly Ala Val Gly Ile Leu Val His Ile
            820                 825                 830

Trp Leu Phe Pro Ile Ile Val Leu Ile Ser Ile Leu Ser Pro Arg Arg
        835                 840                 845
```

<210> 11
<211> 793
<212> PRT
<213> Homo sapiens

<400> 11

```
Ser Gln Asp Ile His Val His Arg Trp Thr Lys Thr Ala Gly Ser Ala
  1               5                  10                  15

Ser Asp Arg Phe Gln Asp Ser Ser Val Phe Asn Glu Thr Leu Arg Ile
             20                  25                  30

Thr Asn Ile Gln Arg His Gln Gly Gly Arg Tyr Tyr Cys Lys Ala Glu
         35                  40                  45

Asn Gly Leu Gly Ser Pro Ala Ile Lys Ser Ile Arg Val Asp Val Tyr
     50                  55                  60

Tyr Leu Asp Asp Pro Val Val Thr Val His Gln Ser Ile Gly Glu Ala
 65                  70                  75                  80

Lys Glu Gln Phe Tyr Tyr Glu Arg Thr Val Phe Leu Arg Cys Val Ala
                85                  90                  95

Asn Ser Asn Pro Pro Val Arg Tyr Ser Trp Arg Arg Gly Gln Glu Val
            100                 105                 110

Leu Leu Gln Gly Ser Asp Lys Gly Val Glu Ile Tyr Glu Pro Phe Phe
        115                 120                 125

Thr Gln Gly Glu Thr Lys Ile Leu Lys Leu Lys Asn Leu Arg Pro Gln
    130                 135                 140

Asp Tyr Ala Asn Tyr Ser Cys Ile Ala Ser Val Arg Asn Val Cys Asn
145                 150                 155                 160

Ile Pro Asp Lys Met Val Ser Phe Arg Leu Ser Asn Lys Thr Ala Ser
                165                 170                 175

Pro Ser Ile Lys Leu Leu Val Asp Asp Pro Ile Val Val Asn Pro Gly
            180                 185                 190

Glu Ala Ile Thr Leu Val Cys Val Thr Thr Gly Gly Glu Pro Ala Pro
        195                 200                 205
```

```
Ser Leu Thr Trp Val Arg Ser Phe Gly Thr Leu Pro Glu Lys Thr Val
    210                 215                 220

Leu Asn Gly Gly Thr Leu Thr Ile Pro Ala Ile Thr Ser Asp Asp Ala
225                 230                 235                 240

Gly Thr Tyr Ser Cys Ile Ala Asn Asn Asn Val Gly Asn Pro Ala Lys
                245                 250                 255

Lys Ser Thr Asn Ile Ile Val Arg Ala Leu Lys Lys Gly Arg Phe Trp
            260                 265                 270

Ile Thr Pro Asp Pro Tyr His Lys Asp Asp Asn Ile Gln Ile Gly Arg
        275                 280                 285

Glu Val Lys Ile Ser Cys Gln Val Glu Ala Val Pro Ser Glu Glu Leu
    290                 295                 300

Thr Phe Ser Trp Phe Lys Asn Gly Arg Pro Leu Arg Ser Ser Glu Arg
305                 310                 315                 320

Met Val Ile Thr Gln Thr Asp Pro Asp Val Ser Pro Gly Thr Thr Asn
                325                 330                 335

Leu Asp Ile Ile Asp Leu Lys Phe Thr Asp Ser Gly Thr Tyr Thr Cys
            340                 345                 350

Val Ala Ser Leu Lys Gly Gly Gly Ile Ser Asp Ile Ser Ile Asp Val
        355                 360                 365

Asn Ile Ser Ser Ser Thr Val Pro Pro Ala Val Glu Pro Ala Phe Leu
    370                 375                 380

Glu Ile Arg Gln Gly Gln Asp Arg Ser Val Thr Met Ser Cys Arg Val
385                 390                 395                 400

Leu Arg Ala Tyr Pro Ile Arg Val Leu Thr Tyr Glu Trp Arg Leu Gly
                405                 410                 415

Asn Lys Leu Leu Arg Thr Gly Gln Phe Asp Ser Gln Glu Tyr Thr Glu
            420                 425                 430

Tyr Ala Val Lys Ser Leu Ser Asn Glu Asn Tyr Gly Val Tyr Asn Cys
        435                 440                 445

Ser Ile Ile Asn Glu Ala Gly Ala Gly Arg Cys Ser Phe Leu Val Thr
    450                 455                 460

Gly Lys Ala Tyr Ala Pro Glu Phe Tyr Tyr Asp Thr Tyr Asn Pro Val
465                 470                 475                 480

Trp Gln Asn Arg His Arg Val Tyr Ser Tyr Ser Leu Gln Trp Thr Gln
                485                 490                 495

Met Asn Pro Asp Ala Val Asp Arg Ile Val Ala Tyr Arg Leu Gly Ile
            500                 505                 510

Arg Gln Ala Gly Gln Gln Arg Trp Trp Glu Gln Glu Ile Lys Ile Tyr
        515                 520                 525
```

```
Gly Asn Ile Gln Lys Gly Glu Leu Ile Thr Tyr Asn Leu Thr Glu Leu
    530                 535                 540

Ile Lys Pro Glu Ala Tyr Glu Val Arg Leu Thr Pro Leu Thr Lys Phe
545                 550                 555                 560

Gly Glu Gly Asp Ser Thr Ile Arg Val Ile Lys Tyr Ser Ala Pro Val
                565                 570                 575

Asn Pro His Leu Arg Glu Phe His Cys Gly Phe Glu Asp Gly Asn Ile
            580                 585                 590

Cys Leu Phe Thr Gln Asp Asp Thr Asp Asn Phe Asp Trp Thr Lys Gln
        595                 600                 605

Ser Thr Ala Thr Arg Asn Thr Lys Tyr Thr Pro Asn Thr Gly Pro Asn
    610                 615                 620

Ala Asp Arg Ser Gly Ser Lys Glu Gly Phe Tyr Met Tyr Ile Glu Thr
625                 630                 635                 640

Ser Arg Pro Arg Leu Glu Gly Glu Lys Ala Arg Leu Leu Ser Pro Val
                645                 650                 655

Phe Ser Ile Ala Pro Lys Asn Pro Tyr Gly Pro Thr Asn Thr Ala Tyr
            660                 665                 670

Cys Phe Ser Phe Phe Tyr His Met Tyr Gly Gln His Ile Gly Val Leu
        675                 680                 685

Asn Val Tyr Leu Arg Leu Lys Gly Gln Thr Thr Ile Glu Asn Pro Leu
    690                 695                 700

Trp Ser Ser Ser Gly Asn Lys Gly Gln Arg Trp Asn Glu Ala His Val
705                 710                 715                 720

Asn Ile Tyr Pro Ile Thr Ser Phe Gln Leu Ile Phe Glu Gly Ile Arg
                725                 730                 735

Gly Pro Gly Ile Glu Gly Asp Ile Ala Ile Asp Asp Val Ser Ile Ala
            740                 745                 750

Glu Gly Glu Cys Ala Lys Gln Asp Leu Ala Thr Lys Asn Ser Val Asp
        755                 760                 765

Gly Ala Val Gly Ile Leu Val His Ile Trp Leu Phe Pro Ile Ile Val
    770                 775                 780

Leu Ile Ser Ile Leu Ser Pro Arg Arg
785                 790
```

<210> 12
<211> 891
<212> PRT
<213> Homo sapiens

```
<400> 12
Ser Gln Asp Ile His Val His Arg Trp Thr Lys Thr Ala Gly Ser Ala
  1               5                  10                  15
```

```
Ser Asp Arg Phe Gln Asp Ser Ser Val Phe Asn Glu Thr Leu Arg Ile
                20                  25                  30

Thr Asn Ile Gln Arg His Gln Gly Gly Arg Tyr Tyr Cys Lys Ala Glu
            35                  40                  45

Asn Gly Leu Gly Ser Pro Ala Ile Lys Ser Ile Arg Val Asp Val Tyr
        50                  55                  60

Tyr Leu Asp Asp Pro Val Val Thr Val His Gln Ser Ile Gly Glu Ala
    65                  70                  75                  80

Lys Glu Gln Phe Tyr Tyr Glu Arg Thr Val Phe Leu Arg Cys Val Ala
                85                  90                  95

Asn Ser Asn Pro Pro Val Arg Tyr Ser Trp Arg Arg Gly Gln Glu Val
            100                 105                 110

Leu Leu Gln Gly Ser Asp Lys Gly Val Glu Ile Tyr Glu Pro Phe Phe
        115                 120                 125

Thr Gln Gly Glu Thr Lys Ile Leu Lys Leu Lys Asn Leu Arg Pro Gln
    130                 135                 140

Asp Tyr Ala Asn Tyr Ser Cys Ile Ala Ser Val Arg Asn Val Cys Asn
145                 150                 155                 160

Ile Pro Asp Lys Met Val Ser Phe Arg Leu Ser Asn Lys Thr Ala Ser
                165                 170                 175

Pro Ser Ile Lys Leu Leu Val Asp Asp Pro Ile Val Val Asn Pro Gly
            180                 185                 190

Glu Ala Ile Thr Leu Val Cys Val Thr Thr Gly Gly Glu Pro Ala Pro
        195                 200                 205

Ser Leu Thr Trp Val Arg Ser Phe Gly Thr Leu Pro Glu Lys Thr Val
    210                 215                 220

Leu Asn Gly Gly Thr Leu Thr Ile Pro Ala Ile Thr Ser Asp Asp Ala
225                 230                 235                 240

Gly Thr Tyr Ser Cys Ile Ala Asn Asn Asn Val Gly Asn Pro Ala Lys
                245                 250                 255

Lys Ser Thr Asn Ile Ile Val Arg Ala Leu Lys Lys Gly Arg Phe Trp
            260                 265                 270

Ile Thr Pro Asp Pro Tyr His Lys Asp Asp Asn Ile Gln Ile Gly Arg
        275                 280                 285

Glu Val Lys Ile Ser Cys Gln Val Glu Ala Val Pro Ser Glu Glu Leu
    290                 295                 300

Thr Phe Ser Trp Phe Lys Asn Gly Arg Pro Leu Arg Ser Ser Glu Arg
305                 310                 315                 320

Met Val Ile Thr Gln Thr Asp Pro Asp Val Ser Pro Gly Thr Thr Asn
                325                 330                 335

Leu Asp Ile Ile Asp Leu Lys Phe Thr Asp Ser Gly Thr Tyr Thr Cys
```

```
            340                 345                 350

Val Ala Ser Leu Lys Gly Gly Gly Ile Ser Asp Ile Ser Ile Asp Val
        355                 360                 365

Asn Ile Ser Ser Ser Thr Val Pro Pro Asn Leu Thr Val Pro Gln Glu
    370                 375                 380

Lys Ser Pro Leu Val Thr Arg Glu Gly Asp Thr Ile Glu Leu Gln Cys
385                 390                 395                 400

Gln Val Thr Gly Lys Pro Lys Pro Ile Ile Leu Trp Ser Arg Ala Asp
                405                 410                 415

Lys Glu Val Ala Met Pro Asp Gly Ser Met Gln Met Glu Ser Tyr Asp
            420                 425                 430

Gly Thr Leu Arg Ile Val Asn Val Ser Arg Glu Met Ser Gly Met Tyr
        435                 440                 445

Arg Cys Gln Thr Ser Gln Tyr Asn Gly Phe Asn Val Lys Pro Arg Glu
    450                 455                 460

Ala Leu Val Gln Leu Ile Val Gln Tyr Pro Pro Ala Val Glu Pro Ala
465                 470                 475                 480

Phe Leu Glu Ile Arg Gln Gly Gln Asp Arg Ser Val Thr Met Ser Cys
                485                 490                 495

Arg Val Leu Arg Ala Tyr Pro Ile Arg Val Leu Thr Tyr Glu Trp Arg
            500                 505                 510

Leu Gly Asn Lys Leu Leu Arg Thr Gly Gln Phe Asp Ser Gln Glu Tyr
        515                 520                 525

Thr Glu Tyr Ala Val Lys Ser Leu Ser Asn Glu Asn Tyr Gly Val Tyr
    530                 535                 540

Asn Cys Ser Ile Ile Asn Glu Ala Gly Ala Gly Arg Cys Ser Phe Leu
545                 550                 555                 560

Val Thr Gly Lys Ala Tyr Ala Pro Glu Phe Tyr Tyr Asp Thr Tyr Asn
                565                 570                 575

Pro Val Trp Gln Asn Arg His Arg Val Tyr Ser Tyr Ser Leu Gln Trp
            580                 585                 590

Thr Gln Met Asn Pro Asp Ala Val Asp Arg Ile Val Ala Tyr Arg Leu
        595                 600                 605

Gly Ile Arg Gln Ala Gly Gln Gln Arg Trp Trp Glu Gln Glu Ile Lys
    610                 615                 620

Ile Asn Gly Asn Ile Gln Lys Gly Glu Leu Ile Thr Tyr Asn Leu Thr
625                 630                 635                 640

Glu Leu Ile Lys Pro Glu Ala Tyr Glu Val Arg Leu Thr Pro Leu Thr
                645                 650                 655

Lys Phe Gly Glu Gly Asp Ser Thr Ile Arg Val Ile Lys Tyr Ser Ala
            660                 665                 670
```

```
Pro Val Asn Pro His Leu Arg Glu Phe His Cys Gly Phe Glu Asp Gly
        675                 680                 685

Asn Ile Cys Leu Phe Thr Gln Asp Asp Thr Asp Asn Phe Asp Trp Thr
    690                 695                 700

Lys Gln Ser Thr Ala Thr Arg Asn Thr Lys Tyr Thr Pro Asn Thr Gly
705                 710                 715                 720

Pro Asn Ala Asp Arg Ser Gly Ser Lys Glu Gly Phe Tyr Met Tyr Ile
            725                 730                 735

Glu Thr Ser Arg Pro Arg Leu Glu Gly Glu Lys Ala Arg Leu Leu Ser
            740                 745                 750

Pro Val Phe Ser Ile Ala Pro Lys Asn Pro Tyr Gly Pro Thr Asn Thr
        755                 760                 765

Ala Tyr Cys Phe Ser Phe Phe Tyr His Met Tyr Gly Gln His Ile Gly
    770                 775                 780

Val Leu Asn Val Tyr Leu Arg Leu Lys Gly Gln Thr Thr Ile Glu Asn
785                 790                 795                 800

Pro Leu Trp Ser Ser Ser Gly Asn Lys Gly Gln Arg Trp Asn Glu Ala
                805                 810                 815

His Val Asn Ile Tyr Pro Ile Thr Ser Phe Gln Leu Ile Phe Glu Gly
            820                 825                 830

Ile Arg Gly Pro Gly Ile Glu Gly Asp Ile Ala Ile Asp Asp Val Ser
        835                 840                 845

Ile Ala Glu Gly Glu Cys Ala Lys Gln Asp Leu Ala Thr Lys Asn Ser
    850                 855                 860

Val Asp Gly Ala Val Gly Ile Leu Val His Ile Trp Leu Phe Pro Ile
865                 870                 875                 880

Ile Val Leu Ile Ser Ile Leu Ser Pro Arg Arg
                885                 890
```

<210> 13
<211> 946
<212> PRT
<213> Homo sapiens

```
<400> 13
Phe Phe Thr Asn Lys Ala Glu Asp Gln Arg Cys Leu Ala Ser Ala Pro
  1               5                  10                  15

Pro Thr Val Arg Ile Val His Ser Gly Leu Ala Cys Asn Ile Glu Glu
             20                  25                  30

Glu Arg Tyr Ser Glu Arg Val Tyr Thr Ile Arg Glu Gly Glu Thr Leu
         35                  40                  45

Glu Leu Thr Cys Leu Val Thr Gly His Pro Arg Pro Gln Ile Arg Trp
     50                  55                  60
```

```
Thr Lys Thr Ala Gly Ser Ala Ser Asp Arg Phe Gln Asp Ser Ser Val
 65                  70                  75                  80

Phe Asn Glu Thr Leu Arg Ile Thr Asn Ile Gln Arg His Gln Gly Gly
                 85                  90                  95

Arg Tyr Tyr Cys Lys Ala Glu Asn Gly Leu Gly Ser Pro Ala Ile Lys
            100                 105                 110

Ser Ile Arg Val Asp Val Tyr Tyr Leu Asp Asp Pro Val Val Thr Val
        115                 120                 125

His Gln Ser Ile Gly Glu Ala Lys Glu Gln Phe Tyr Tyr Glu Arg Thr
    130                 135                 140

Val Phe Leu Arg Cys Val Ala Asn Ser Asn Pro Pro Val Arg Tyr Ser
145                 150                 155                 160

Trp Arg Arg Gly Gln Glu Val Leu Leu Gln Gly Ser Asp Lys Gly Val
                165                 170                 175

Glu Ile Tyr Glu Pro Phe Phe Thr Gln Gly Glu Thr Lys Ile Leu Lys
            180                 185                 190

Leu Lys Asn Leu Arg Pro Gln Asp Tyr Ala Asn Tyr Ser Cys Ile Ala
        195                 200                 205

Ser Val Arg Asn Val Cys Asn Ile Pro Asp Lys Met Val Ser Phe Arg
    210                 215                 220

Leu Ser Asn Lys Thr Ala Ser Pro Ser Ile Lys Leu Leu Val Asp Asp
225                 230                 235                 240

Pro Ile Val Val Asn Pro Gly Glu Ala Ile Thr Leu Val Cys Val Thr
                245                 250                 255

Thr Gly Gly Glu Pro Ala Pro Ser Leu Thr Trp Val Arg Ser Phe Gly
            260                 265                 270

Thr Leu Pro Glu Lys Thr Val Leu Asn Gly Gly Thr Leu Thr Ile Pro
        275                 280                 285

Ala Ile Thr Ser Asp Asp Ala Gly Thr Tyr Ser Cys Ile Ala Asn Asn
    290                 295                 300

Asn Val Gly Asn Pro Ala Lys Lys Ser Thr Asn Ile Ile Val Arg Ala
305                 310                 315                 320

Leu Lys Lys Gly Arg Phe Trp Ile Thr Pro Asp Pro Tyr His Lys Asp
                325                 330                 335

Asp Asn Ile Gln Ile Gly Arg Glu Val Lys Ile Ser Cys Gln Val Glu
            340                 345                 350

Ala Val Pro Ser Glu Glu Leu Thr Phe Ser Trp Phe Lys Asn Gly Arg
        355                 360                 365

Pro Leu Arg Ser Ser Glu Arg Met Val Ile Thr Gln Thr Asp Pro Asp
    370                 375                 380
```

```
Val Ser Pro Gly Thr Thr Asn Leu Asp Ile Ile Asp Leu Lys Phe Thr
385                 390                 395                 400

Asp Ser Gly Thr Tyr Thr Cys Val Ala Ser Leu Lys Gly Gly Gly Ile
                405                 410                 415

Ser Asp Ile Ser Ile Asp Val Asn Ile Ser Ser Ser Thr Val Pro Pro
            420                 425                 430

Asn Leu Thr Val Pro Gln Glu Lys Ser Pro Leu Val Thr Arg Glu Gly
        435                 440                 445

Asp Thr Ile Glu Leu Gln Cys Gln Val Thr Gly Lys Pro Lys Pro Ile
    450                 455                 460

Ile Leu Trp Ser Arg Ala Asp Lys Glu Val Ala Met Pro Asp Gly Ser
465                 470                 475                 480

Met Gln Met Glu Ser Tyr Asp Gly Thr Leu Arg Ile Val Asn Val Ser
                485                 490                 495

Arg Glu Met Ser Gly Met Tyr Arg Cys Gln Thr Ser Gln Tyr Asn Gly
            500                 505                 510

Phe Asn Val Lys Pro Arg Glu Ala Leu Val Gln Leu Ile Val Gln Tyr
        515                 520                 525

Pro Pro Ala Val Glu Pro Ala Phe Leu Glu Ile Arg Gln Gly Gln Asp
    530                 535                 540

Arg Ser Val Thr Met Ser Cys Arg Val Leu Arg Ala Tyr Pro Ile Arg
545                 550                 555                 560

Val Leu Thr Tyr Glu Trp Arg Leu Gly Asn Lys Leu Leu Arg Thr Gly
                565                 570                 575

Gln Phe Asp Ser Gln Glu Tyr Thr Glu Tyr Ala Val Lys Ser Leu Ser
            580                 585                 590

Asn Glu Asn Tyr Gly Val Tyr Asn Cys Ser Ile Ile Asn Glu Ala Gly
        595                 600                 605

Ala Gly Arg Cys Ser Phe Leu Val Thr Gly Lys Ala Tyr Ala Pro Glu
    610                 615                 620

Phe Tyr Tyr Asp Thr Tyr Asn Pro Val Trp Gln Asn Arg His Arg Val
625                 630                 635                 640

Tyr Ser Tyr Ser Leu Gln Trp Thr Gln Met Asn Pro Asp Ala Val Asp
                645                 650                 655

Arg Ile Val Ala Tyr Arg Leu Gly Ile Arg Gln Ala Gly Gln Gln Arg
            660                 665                 670

Trp Trp Glu Gln Glu Ile Lys Ile Asn Gly Asn Ile Gln Lys Gly Glu
        675                 680                 685

Leu Ile Thr Tyr Asn Leu Thr Glu Leu Ile Lys Pro Glu Ala Tyr Glu
    690                 695                 700

Val Arg Leu Thr Pro Leu Thr Lys Phe Gly Glu Gly Asp Ser Thr Ile
```

705 710 715 720

Arg Val Ile Lys Tyr Ser Ala Pro Val Asn Pro His Leu Arg Glu Phe
725 730 735

His Cys Gly Phe Glu Asp Gly Asn Ile Cys Leu Phe Thr Gln Asp Asp
740 745 750

Thr Asp Asn Phe Asp Trp Thr Lys Gln Ser Thr Ala Thr Arg Asn Thr
755 760 765

Lys Tyr Thr Pro Asn Thr Gly Pro Asn Ala Asp Arg Ser Gly Ser Lys
770 775 780

Glu Gly Phe Tyr Met Tyr Ile Glu Thr Ser Arg Pro Arg Leu Glu Gly
785 790 795 800

Glu Lys Ala Arg Leu Leu Ser Pro Val Phe Ser Ile Ala Pro Lys Asn
805 810 815

Pro Tyr Gly Pro Thr Asn Thr Ala Tyr Cys Phe Ser Phe Phe Tyr His
820 825 830

Met Tyr Gly Gln His Ile Gly Val Leu Asn Val Tyr Leu Arg Leu Lys
835 840 845

Gly Gln Thr Thr Ile Glu Asn Pro Leu Trp Ser Ser Ser Gly Asn Lys
850 855 860

Gly Gln Arg Trp Asn Glu Ala His Val Asn Ile Tyr Pro Ile Thr Ser
865 870 875 880

Phe Gln Leu Ile Phe Glu Gly Ile Arg Gly Pro Gly Ile Glu Gly Asp
885 890 895

Ile Ala Ile Asp Asp Val Ser Ile Ala Glu Gly Glu Cys Ala Lys Gln
900 905 910

Asp Leu Ala Thr Lys Asn Ser Val Asp Gly Ala Val Gly Ile Leu Val
915 920 925

His Ile Trp Leu Phe Pro Ile Ile Val Leu Ile Ser Ile Leu Ser Pro
930 935 940

Arg Arg
945

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2001 (26.04.2001)

PCT

(10) International Publication Number
WO 01/29215 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/705, 16/28, A61K 38/17, 31/70, 39/395, C12Q 1/68, G01N 33/53

(21) International Application Number: PCT/IL00/00664

(22) International Filing Date: 19 October 2000 (19.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
132459 19 October 1999 (19.10.1999) IL
134932 7 March 2000 (07.03.2000) IL

(71) Applicant *(for all designated States except US)*: **COMPUGEN LTD.** [IL/IL]; Pinchas Rosen St. 72, 69512 Tel Aviv (IL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SAVITZKY, Kinneret** [IL/IL]; Metodela St. 44, 69548 Tel Aviv (IL). **BERNSTEIN, Jeanne** [IL/IL]; Harimon St. 23, 40300 Kfar Yona (IL). **DAVID, Anat** [IL/IL]; Habrosh St. 80, 30500 Binyamina (IL).

(74) Agent: **REINHOLD COHN AND PARTNERS**; P.O.B. 4060, 61040 Tel Aviv (IL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

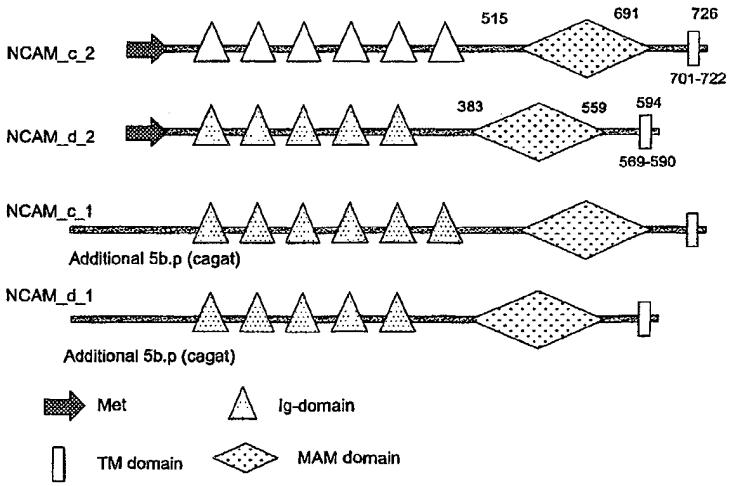
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
27 December 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: N-CAM HOMOLOG AND SPLICE VARIANTS THEREOF



(57) Abstract: The invention concerns several nucleic acid and amino acid sequences which are homologs of the neuronal cell adhesion molecule. These sequences are splice variants of each other.

WO 01/29215 A3

INTERNATIONAL SEARCH REPORT

In' ational Application No
PCT/IL 00/00664

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C07K16/28 A61K38/17 A61K31/70
A61K39/395 C12Q1/68 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A61K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! Acc.no. AI859192, 22 July 1999 (1999-07-22) STRAUSBERG, R. ET AL.: "w166h10.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2429923 3' similar to SW:THIB_XENLA Q91641 THYROID HORMONE-INDUCED PROTEIN B PRECURSOR." XP002173300 Human EST with 99.1% identity with seq.ID.1 between nt.2516-2961. the whole document --- -/--	1-4

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 July 2001	07/08/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Smalt, R

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IL 00/00664

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	SAUGIER-VEBER PASCALE ET AL: "Identification of novel L1CAM mutations using fluorescence-assisted mismatch analysis." HUMAN MUTATION, vol. 12, no. 4, 1998, pages 259-266, XP000990828 ISSN: 1059-7794 the whole document	
A	US 5 840 689 A (DANILOFF JOANNE K) 24 November 1998 (1998-11-24) the whole document	
A	WO 96 04396 A (SYSTEMIX INC) 15 February 1996 (1996-02-15) the whole document	
P,X	DATABASE EMBL 'Online! Acc.no. AB047834, 31 August 2000 (2000-08-31) OSADA, N ET AL.: "Macaca fascicularis brain cDNA, clone:QccE-16296." XP002173301 98.4% identity with seq.ID.1 between nt. 1488-3065; 99.77% identity between amino acid residues 499-946. the whole document	1-5,7, 9-11

2

Continuation of Box I.2

Present claims 22 and 23 relate to an activator and a deactivator, respectively, of the protein of the invention.
The application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compounds by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, namely those parts relating to the (de)activators which can be defined by the protein of the invention, namely antibodies directed to said protein.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/IL 00/00664

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5840689 A	24-11-1998	NONE	
WO 9604396 A	15-02-1996	US 5665590 A	09-09-1997
		AU 3206095 A	04-03-1996